

Full Duplex Assisted Inter-cell Interference Cancellation in Heterogeneous Networks

Shengqian Han, *Member, IEEE*, Chenyang Yang, *Senior Member, IEEE*, and Pan Chen

Abstract—The paper studies the suppression of cross-tier inter-cell interference (ICI) generated by a macro base station (MBS) to pico user equipments (PUEs) in heterogeneous networks (HetNets). Different from existing ICI avoidance schemes such as enhanced ICI cancellation (eICIC) and coordinated beamforming, which generally operate at the MBS, we propose a full duplex (FD) assisted ICI cancellation (fICIC) scheme, which can operate at each pico BS (PBS) individually and is transparent to the MBS. The basic idea of the fICIC is to apply FD technique at the PBS such that the PBS can send the desired signals and forward the listened cross-tier ICI simultaneously to PUEs. We first consider the narrowband single-user case, where the MBS serves a single macro UE and each PBS serves a single PUE. We obtain the closed-form solution of the optimal fICIC scheme, and analyze its asymptotical performance in ICI-dominated scenario. We then investigate the general narrowband multi-user case, where both MBS and PBSs serve multiple UEs. We devise a low-complexity algorithm to optimize the fICIC aimed at maximizing the downlink sum rate of the PUEs subject to user fairness constraint. Finally, the generalization of the fICIC to wideband systems is investigated. Simulations validate the analytical results and demonstrate the advantages of the fICIC on mitigating cross-tier ICI.

Index Terms—Inter-cell interference cancellation, Full duplex, Heterogeneous networks, eICIC, CoMP, OFDM.

I. INTRODUCTION

Cellular systems are evolving toward heterogeneous networks (HetNets) with universal frequency reuse in order to support the galloping demand of mobile wireless services [1]. By deploying low-power nodes such as micro, pico, or femto base stations (BSs) within the coverage of traditional macro BSs (MBSs), HetNets bring the network close to the user equipments (UEs) to obtain the “cell-splitting” gain, and therefore improve the area spectral efficiency. For simplicity, we refer to the low-power nodes as pico BSs (PBSs) hereinafter.

In practice, however, straightforwardly deploying PBSs in the coverage of MBSs cannot effectively realize the promised benefits of HetNets because the large difference of the two types of BSs in transmit power makes pico UEs (PUEs) suffer from severe cross-tier inter-cell interference (ICI) generated by the MBS. Efficient ICI cancellation (ICIC) mechanisms are therefore critical to support the HetNet deployment [1].

To mitigate the ICI, enhanced ICIC (eICIC) techniques have been developed in Release 10 of Long-term Evolution (LTE)

[1]. In the time-domain eICIC, the MBS remains silent in the so-called almost blank subframes (ABS), during which the cell-edge PUEs are served without interference. In the frequency-domain eICIC, the MBS and PBSs schedule UEs in orthogonal frequency resources to avoid the ICI. The eICIC methods are of low complexity and easy to implement, but they limit the performance of both macro UEs (MUEs) and PUEs since the UEs can be only served in partial time-frequency resources. Coordinated multi-point (CoMP) transmission is another promising technique for cross-tier ICI suppression, which exploits the antenna resources of the MBS [2]. Considering the fact that providing high-performance backhaul from all PBSs to the core network may be cost prohibitive, coordinated beamforming (CB) based CoMP transmission, requiring no data sharing among the BSs, has received wide attention. CoMP-CB can be considered as a sort of spatial-domain eICIC, which enables the PUEs to be served with the whole time-frequency resources. However, the performance of CoMP-CB is limited by the number of antennas at the MBS, which is usually not acceptable when the PBSs are densely deployed in the coverage of the MBS.

Different from eICIC and CoMP-CB, both of which control the transmission of the MBS in time, frequency or spatial domain in order to generate an ICI-free environment for the transmission between PBSs and PUEs, in this paper we strive to study the cross-tier ICI suppression scheme without the participation of MBSs, which therefore will not consume the resources of the MBS. Specifically, we consider using full duplex (FD) technique to HetNets. FD communication was long believed impossible in wireless system design due to the severe self-interference within the same transceiver. However, the belief has been overturned recently with the tremendous progress in self-interference cancellation [3, 4], where the plausibility of FD technique for short-range point-to-point communications was approved. FD techniques have been applied to provide bi-directional communications over the same time and frequency resources, and exhibited noticeable spectral efficiency gains over half-duplex (HD) schemes [5, 6]. FD techniques are also applied in relay systems to improve service coverage, where the usage of FD avoids the waste of resources as in HD relay systems [7, 8]. To the best of our knowledge, leveraging FD in HetNets to suppress the cross-tier ICI has not been addressed in the literature.

In this paper, we consider a HetNet consisting of a MBS and multiple PBSs. We investigate the application of FD technique to mitigate the ICI generated by the MBS to the PUEs. The main contributions are summarized as follows:

- We propose a novel ICI cancellation scheme for HetNets,

This work was supported in part by the National Natural Science Foundation of China (No. 61301084) and by the Fundamental Research Funds for the Central Universities.

S. Han, C. Yang, and P. Chen are with the School of Electronics and Information Engineering, Beihang University, Beijing, China (e-mail: {sqhan, cyyang, chenpan}@buaa.edu.cn).

TABLE I
LIST OF MAJOR SYMBOLS

B, K_M, K_P	Numbers of PBSs, MUEs, and PUEs per pico cell
M, N_t	Numbers of transmit antennas at the MBS and each PBS
N_r, N_t	Numbers of FD and HD receive antennas and HD
$\mathbf{h}_{Pk}, \mathbf{H}_{PP}$	Channel from PBS to PUE _k and self-interference channel at the FD PBS
$\mathbf{H}_{MP}, \mathbf{h}_{Mk}$	Equivalent channels from MBS to PBS and PUE _k
$\mathbf{h}_{MP}, \mathbf{h}_{Mk}$	Equivalent channels from MBS to PBS and PUE _k in single-user case
$\bar{\mathbf{g}}_{MPn}, \bar{\mathbf{g}}_{Mkn}$	Equivalent frequency-domain channels from MBS to PBS and PUE _k
$\mathbf{y}'_P, \mathbf{y}_P$	Receive signals of FD PBS before and after self-interference cancellation
$\mathbf{z}_x, \mathbf{z}_y$	Transmitter and receiver distortions from hardware impairments
\mathbf{W}_f, L	Forwarding precoder at PBS and its order in wideband systems
\mathbf{w}_f	Vectorization of \mathbf{W}_f^H
$\mathbf{w}_{d,k}$	Precoder at PBS for desired signals of PUE _k
$\mathbf{W}_{fn}, \bar{\mathbf{w}}_{dn}$	Frequency-domain precoders for forwarded ICI and desired signals
P_{out}, P_0	Total and maximal transmit power of PBS
$P_{out,n}$	Transmit power on the n -th subcarrier of PBS

named FD assisted ICI cancellation (fICIC), where the PBSs are supposed to have the FD capability of transmitting and receiving simultaneously. The basic idea of the proposed fICIC is to let the FD PBSs forward the listened interference from the MBS to the PUEs to neutralize the ICI, while at the same time sending the desired signals. Since the ICI is mitigated at PBSs now, the proposed fICIC is transparent to the MBS in the sense that no changes are needed for the transmission of the MBS. Note that a similar usage of the FD technique was presented in [9] for the cooperative cognitive network, where the FD secondary BS forwards the listened primary signals in order to increase the primary spectrum accessing opportunities, which leads to completely different problem and strategy design from ours.

- In narrowband single-user case, where the MBS serves a single MUE and each PBS serves a single PUE, we first find the explicit expressions of the optimal fICIC precoders, which maximize the signal-to-interference-plus-noise ratio (SINR) of the PUE in each pico cell. We then analyze the asymptotical performance of the fICIC in ICI-dominated scenario. The results show that under perfect self-interference cancellation for FD, the fICIC can thoroughly eliminate weak ICI, while when the ICI is very strong or the residual self-interference is very large, the fICIC will reduce to the HD scheme.
- In narrowband multi-user case, where the MBS serves multiple MUEs and each PBS serves multiple PUEs, we propose a low-complexity algorithm to optimize the fICIC scheme, aimed at maximizing the downlink sum rate of each pico cell under the fairness constraint over the PUEs. Simulations validate the analytical results, and demonstrate a significant performance gain of the fICIC over the HD scheme as well as evident benefits of combining the fICIC with existing eICIC and CoMP techniques.
- We generalize the narrowband fICIC to orthogonal

frequency-division multiplexing (OFDM) systems, where the optimization problem for the fICIC precoder design aimed at maximizing sum rate over multiple subcarriers is obtained, which can be solved with a gradient based method. Simulations shows the advantages of the fICIC on mitigating wideband ICI.

Notations: $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote transpose, complex conjugate, and conjugate transpose, respectively. $\mathbf{X} \succeq \mathbf{0}$ represents that matrix \mathbf{X} is positive semi-definite, and $\mathbf{X}^{\frac{1}{2}}$ denotes hermitian square root of \mathbf{X} . $\text{tr}(\cdot)$ denotes matrix trace, $\text{rank}(\cdot)$ denotes matrix rank, $\mathcal{E}\{\cdot\}$ denotes expectation operator, $\text{vec}(\cdot)$ denotes vectorization operator, \otimes denotes Kronecker product, \odot denotes convolution product, and $\|\cdot\|$ denotes Euclidian norm. $\text{diag}(\mathbf{X})$ denotes the diagonal matrix with the same diagonal elements as \mathbf{X} . \mathbf{I}_N , $\mathbf{0}_N$ and $\bar{\mathbf{0}}_N$ denote $N \times N$ identity and zero matrices, and $N \times 1$ zero vector, respectively. The major symbols used in the paper are listed in Table I.

II. SYSTEM MODEL

We first consider the downlink transmission of a narrowband time division duplex (TDD) HetNet, and then generalize the model to wideband systems. Suppose that the HetNet consists of one MBS and B PBSs, where the MBS serves K_M single-antenna MUEs and each PBS serves K_P single-antenna PUEs. Assume that the MUEs experience negligible interference from PBSs due to the coverage range expansion (CRE) of pico cells, and the pico cells are geographically separated so that each PUE receives much weaker interference from interfering PBSs compared to the interference generated by the MBS, which is treated as noise in the paper. Therefore, we focus on the suppression of the cross-tier ICI generated by the MBS to PUEs, which is commonly recognized as a bottleneck to improve the spectral efficiency in real-world HetNets [1].

We consider applying FD technique at each PBS in the downlink transmission, with which the PBS can send the desired signals and forward the listened cross-tier ICI simultaneously. The structure of the FD PBS transceiver is illustrated in Fig. 1(a), where uplink and downlink data flows are indicated by dashed and solid lines, respectively. The FD PBS is comprised of traditional baseband (BB) and radio frequency (RF) modules in HD transceiver, taking charge of the transmission and reception of desired signals of PUEs, as well as additional FD modules, dedicated for self-interference cancellation and ICI suppression. The antennas at the PBS can be divided into two parts as shown in Fig. 1(a), where the “FD receive antennas” are only used to receive the ICI from the MBS in the downlink, while the “transmit antennas” and the “HD receive antennas” share the same antennas in a TDD manner, which are used to send and receive the signals of PUEs in downlink and uplink, respectively. Therefore, the uplink and downlink channel reciprocity holds between the PBS and the PUE.

Since the proposed fICIC scheme will not affect the performance of MUEs and other-cell PUEs, in the sequel we only consider a reference PBS and focus on the performance of the PUEs served by the reference PBS. The resulting interference

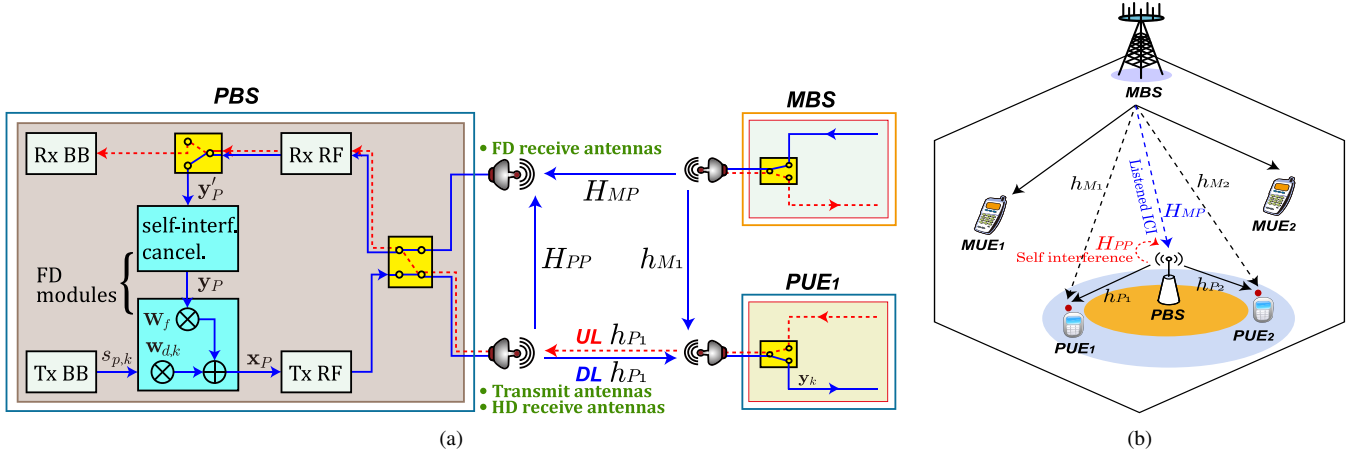


Fig. 1. (a) System model of the TDD HetNet, where the transceiver structure of the FD PBS is illustrated in the left part. (b) Illustration of the considered HetNet layout, where the MBS serves two MUEs and the reference PBS serves two PUEs.

environment is demonstrated in Fig. 1(b). Suppose that the MBS has M transmit antennas, $M \geq K_M$, and the FD PBS has N_t transmit antennas, N_t HD receive antennas, and N_r FD receive antennas, $N_t \geq K_P$. Let $\mathbf{h}_{Mk} \in \mathbb{C}^{M \times 1}$ and $\mathbf{h}_{Pk} \in \mathbb{C}^{N_t \times 1}$ denote the channels from the MBS and the PBS to the k -th PUE (denoted by PUE _{k}), $\mathbf{H}_{MP} \in \mathbb{C}^{M \times N_r}$ denote the channel from the MBS to the PBS, and $\mathbf{H}_{PP} \in \mathbb{C}^{N_t \times N_r}$ denote the self-interference channel of the FD PBS.

A. Signals of the FD PBS

In the downlink the FD PBS can transmit and receive signals simultaneously. We use the index t to denote time instant. To reflect the impact of hardware impairments of transmitter chains on self-interference cancellation, we can express the receive signal at the PBS before self-interference cancellation based on [8, 9] as

$$\bar{\mathbf{y}}_p[t] = \mathbf{H}_{MP}^H \mathbf{W}_M \mathbf{s}_M[t] + \mathbf{H}_{PP}^H (\mathbf{x}_p[t] + \mathbf{z}_x[t]) + \mathbf{n}_p[t], \quad (1)$$

where $\mathbf{W}_M \in \mathbb{C}^{M \times K_M}$ is the precoding matrix at the MBS for sending the signals $\mathbf{s}_M \sim \mathcal{CN}(\bar{\mathbf{0}}_{K_M}, \mathbf{I}_{K_M})$ to the MUEs, the term $\mathbf{H}_{PP}^H (\mathbf{x}_p[t] + \mathbf{z}_x[t])$ is the self-interference, $\mathbf{x}_p \in \mathbb{C}^{N_t \times 1}$ is the transmit signal vector of the FD PBS, $\mathbf{z}_x \sim \mathcal{CN}(\bar{\mathbf{0}}_{N_t}, \mu_x \text{diag}(\Phi_x))$ is the transmitter distortion with Φ_x denoting the covariance matrix of \mathbf{x}_p , $\mu_x \ll 1$ is a scaling constant, which reflects the combined effects of additive power-amplifier noise, non-linearities in digital-to-analog converter and power amplifier, I/Q imbalance, and oscillator phase noise, and $\mathbf{n}_p \sim \mathcal{CN}(\bar{\mathbf{0}}_{N_r}, \sigma_n^2 \mathbf{I}_{N_r})$ is the additive white Gaussian noise (AWGN), which takes into account both thermal noises and the ICI from other PBSs.

Further considering the hardware impairments of receiver chains based on [8, 9], the distorted receive signal can be expressed as

$$\mathbf{y}'_p[t] = \bar{\mathbf{y}}_p[t] + \mathbf{z}_y[t], \quad (2)$$

where $\mathbf{z}_y \sim \mathcal{CN}(\bar{\mathbf{0}}_{N_r}, \mu_y \text{diag}(\Phi_y))$ is the additive distortion caused by adaptive gain control noise, non-linearities in analog-to-digital converter and gain control, I/Q imbalance, and oscillator phase noise in receiver chains, $\mu_y \ll 1$ is

a scaling constant, and Φ_y is the covariance matrix of the undistorted receive signal $\bar{\mathbf{y}}_p$, which can be obtained from (1) as

$$\begin{aligned} \Phi_y &= \mathcal{E}_{\mathbf{s}_M, \mathbf{x}_p, \mathbf{z}_x, \mathbf{n}_p} \{\bar{\mathbf{y}}_p[t] \bar{\mathbf{y}}_p^H[t]\} = \mathbf{H}_{MP}^H \mathbf{W}_M \mathbf{W}_M^H \mathbf{H}_{MP} \\ &+ \mathbf{H}_{PP}^H (\Phi_x + \mu_x \text{diag}(\Phi_x)) \mathbf{H}_{PP} + \sigma_n^2 \mathbf{I}_{N_r}. \end{aligned} \quad (3)$$

Since the transmit signal \mathbf{x}_p is known at the FD PBS, the self-interference $\mathbf{H}_{PP}^H \mathbf{x}_p$ in (1) can be cancelled if the self-interference channel \mathbf{H}_{PP} is estimated. During a dedicated training phase, the orthogonal pilot signals $\sqrt{P_{tr}} \mathbf{C}$ are transmitted for self-interference channel estimation, where P_{tr} is the transmit power of pilot signals, and $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_{N_t}] \in \mathbb{C}^{N_t \times N_t}$ is unitary with $\mathbf{C} \mathbf{C}^H = \mathbf{I}_{N_t}$. Then, similar to (1) and (2), we can obtain the receive pilot signals with hardware impairments of both transmitter and receiver chains as

$$\mathbf{Y}[t] = \tilde{\mathbf{Y}}[t] + \tilde{\mathbf{Z}}_y[t] \triangleq \mathbf{H}_{PP}^H (\sqrt{P_{tr}} \mathbf{C} + \tilde{\mathbf{Z}}_x[t]) + \tilde{\mathbf{N}}[t] + \tilde{\mathbf{Z}}_y[t], \quad (4)$$

where $\tilde{\mathbf{Y}} = [\tilde{\mathbf{y}}_1, \dots, \tilde{\mathbf{y}}_{N_t}] \in \mathbb{C}^{N_r \times N_t}$ denotes the undistorted receive signal, $\tilde{\mathbf{Z}}_y = [\tilde{\mathbf{z}}_{y,1}, \dots, \tilde{\mathbf{z}}_{y,N_t}] \in \mathbb{C}^{N_r \times N_t}$ is the additive distortion of receive signals with $\tilde{\mathbf{z}}_{y,i} \sim \mathcal{CN}(\bar{\mathbf{0}}_{N_r}, \mu_y \text{diag}(\Phi_{y,i}))$, $\Phi_{y,i}$ is the covariance matrix of $\tilde{\mathbf{y}}_i$, $\tilde{\mathbf{Z}}_x = [\tilde{\mathbf{z}}_{x,1}, \dots, \tilde{\mathbf{z}}_{x,N_t}] \in \mathbb{C}^{N_t \times N_t}$ is the transmitter distortion with $\tilde{\mathbf{z}}_{x,i} \sim \mathcal{CN}(\bar{\mathbf{0}}_{N_t}, \mu_x P_{tr} \text{diag}(\mathbf{c}_i \mathbf{c}_i^H))$, and $\tilde{\mathbf{N}} \in \mathbb{C}^{N_r \times N_t}$ is the AWGN whose columns follow the distribution $\mathcal{CN}(\bar{\mathbf{0}}_{N_r}, \sigma_n^2 \mathbf{I}_{N_r})$. It can be obtained from (4) that $\tilde{\Phi}_{y,i} = P_{tr} \mathbf{H}_{PP}^H (\mathbf{c}_i \mathbf{c}_i^H + \mu_x \text{diag}(\mathbf{c}_i \mathbf{c}_i^H)) \mathbf{H}_{PP} + \sigma_n^2 \mathbf{I}_{N_r}$.

With least-squares channel estimator, we can estimate the self-interference channel as

$$\begin{aligned} \hat{\mathbf{H}}_{PP} &= \frac{1}{\sqrt{P_{tr}}} \mathbf{C} \mathbf{Y}^H[t] = \mathbf{H}_{PP} + \frac{1}{\sqrt{P_{tr}}} (\mathbf{C} \tilde{\mathbf{Z}}_x^H[t] \mathbf{H}_{PP} + \\ &\mathbf{C} \tilde{\mathbf{N}}^H[t] + \mathbf{C} \tilde{\mathbf{Z}}_y^H[t]) \triangleq \mathbf{H}_{PP} + \mathbf{E}_{PP}[t], \end{aligned} \quad (5)$$

where $\mathbf{E}_{PP} = [\mathbf{e}_{PP,1}, \dots, \mathbf{e}_{PP,N_t}]^H \in \mathbb{C}^{N_t \times N_r}$ denotes the channel estimation errors.

Denoting $\mathbf{c}_i = [c_{i1}, \dots, c_{iN_t}]^T$, we can express $\mathbf{e}_{PP,i}$ as

$$\mathbf{e}_{PP,i} = \frac{1}{\sqrt{P_{tr}}} \left(\mathbf{H}_{PP}^H \sum_{j=1}^{N_t} c_{ij} \tilde{\mathbf{z}}_{x,j} + \tilde{\mathbf{N}} \mathbf{c}_i + \sum_{j=1}^{N_t} c_{ij} \tilde{\mathbf{z}}_{y,j} \right), \quad (6)$$

from which we can obtain that $\mathbf{e}_{PP,i}$ follows the distribution $\mathcal{CN}(\mathbf{0}_{N_r}, \tilde{\Phi}_{e,i})$ with

$$\tilde{\Phi}_{e,i} = \mathbf{H}_{PP}^H \sum_{j=1}^{N_t} |c_{ij}|^2 \mu_x \text{diag}(\mathbf{c}_j \mathbf{c}_j^H) \mathbf{H}_{PP} + \frac{(1+\mu_y)\sigma_n^2}{P_{tr}} \mathbf{I}_{N_r} + \sum_{j=1}^{N_t} |c_{ij}|^2 \mu_y \text{diag}(\mathbf{H}_{PP}^H (\mathbf{c}_j \mathbf{c}_j^H + \mu_x \text{diag}(\mathbf{c}_j \mathbf{c}_j^H)) \mathbf{H}_{PP}). \quad (7)$$

With $\hat{\mathbf{H}}_{PP}$, the receive signal at the PBS after self-interference cancellation is

$$\mathbf{y}_p[t] = \mathbf{y}'_p[t] - \hat{\mathbf{H}}_{PP}^H \mathbf{x}_p[t] \triangleq \bar{\mathbf{H}}_{MP}^H \mathbf{s}_M[t] - \mathbf{E}_{PP}^H \mathbf{x}_p[t] + \mathbf{H}_{PP}^H \mathbf{z}_x[t] + \mathbf{n}_p[t] + \mathbf{z}_y[t], \quad (8)$$

where $\bar{\mathbf{H}}_{MP} \triangleq \mathbf{W}_M^H \mathbf{H}_{MP}$ is the equivalent channel from the MBS to the FD PBS.

The FD PBS then transmits the desired signals of the PUEs together with the self-interference cancelled receive signals \mathbf{y}_p . The combined transmit signal of the PBS can be expressed as

$$\mathbf{x}_p[t] = \mathbf{W}_f \mathbf{y}_p[t - \tau] + \sum_{k=1}^{K_P} \mathbf{w}_{d,k} s_{p,k}[t], \quad (9)$$

where τ is the processing delay introduced by the FD modules, $\mathbf{W}_f \in \mathbb{C}^{N_t \times N_r}$ is the precoding matrix for the forwarded signals, $\mathbf{w}_{d,k} \in \mathbb{C}^{N_t \times 1}$ is the precoding vector for the desired signal, $s_{p,k}$, of PUE $_k$, and $s_{p,k} \sim \mathcal{CN}(0, 1)$.

In (9), the receive signal \mathbf{y}_p by N_r additional FD receive antennas of the PBS are forwarded via \mathbf{W}_f . The forwarded signal will occupy a part of transmit power of the PBS, which leads to the reduction of the transmit power for desired signals. As will be clear later, however, the forwarded signal can be used to mitigate the cross-tier ICI at PUEs efficiently with the optimized \mathbf{W}_f , and hence result in the improvement of PUEs' data rate.

With (8) and (9) we can calculate the transmit power of the FD PBS as

$$P_{out} = \text{tr}(\Phi_x) = \text{tr}(\mathcal{E}_{\mathbf{s}_M, s_{p,k}, \mathbf{n}_p, \mathbf{z}_x, \mathbf{z}_y, \mathbf{E}_{PP}, \mathbf{H}_{PP}} \{\mathbf{x}_p[t] \mathbf{x}_p^H[t]\}), \quad (10)$$

where the expectations are taken over data \mathbf{s}_M and $s_{p,k}$, noises \mathbf{n}_p , transmitter and receiver distortions \mathbf{z}_x and \mathbf{z}_y , channel estimation errors \mathbf{E}_{PP} , and self-interference channel \mathbf{H}_{PP} , respectively.¹

Assume that \mathbf{H}_{PP} follows Rayleigh distribution, i.e., $\text{vec}(\mathbf{H}_{PP}) \sim \mathcal{CN}(\mathbf{0}_{N_r N_t}, \bar{\alpha}_{PP} \mathbf{I}_{N_r N_t})$ with $\bar{\alpha}_{PP}$ denoting the average channel gain, which is reasonable because the transmit antennas and the FD receive antennas can be well isolated in the considered scenario as will be detailed in Section VI-C. Then, we show in Appendix A that the transmit power P_{out}

¹In order to highlight the benefits of the fICIC scheme via forwarding the listened ICI, we restrict ourselves to the case where the precoders \mathbf{W}_f and $\{\mathbf{w}_{d,k}\}$ are designed only for transmitting the listened ICI and desired signals but not for spatial-domain self-interference cancellation (i.e., design \mathbf{W}_f and $\{\mathbf{w}_{d,k}\}$ to pre-null self-interference). Towards this end, we take the expectation over \mathbf{H}_{PP} in (10), such that the precoders are independent of the instantaneous self-interference channel \mathbf{H}_{PP} .

can be expressed as

$$P_{out} \approx \text{tr}(\mathbf{W}_f \bar{\mathbf{H}}_{MP}^H \bar{\mathbf{H}}_{MP} \mathbf{W}_f^H) + \sum_{k=1}^{K_P} \|\mathbf{w}_{d,k}\|^2 + (\sigma_n^2 + \sigma_e^2 P_{out}) \text{tr}(\mathbf{W}_f \mathbf{W}_f^H), \quad (11)$$

where the approximation follows from $\mu_x \ll 1$ and $\mu_y \ll 1$ as in [8], and $\sigma_e^2 = \frac{\sigma_n^2}{P_{tr}} + 2\bar{\alpha}_{PP}(\mu_x + \mu_y)$ reflects the residual self-interference, in which the term $\frac{\sigma_n^2}{P_{tr}}$ comes from imperfect channel estimation and the term $2\bar{\alpha}_{PP}(\mu_x + \mu_y)$ comes from hardware impairments.

From the equation with respect to P_{out} given in (11), we can obtain the transmit power of the FD PBS as

$$P_{out} = \frac{\text{tr}(\bar{\mathbf{H}}_{MP} \mathbf{W}_f^H \mathbf{W}_f \bar{\mathbf{H}}_{MP}^H) + \sigma_n^2 \text{tr}(\mathbf{W}_f^H \mathbf{W}_f) + \sum_{k=1}^{K_P} \|\mathbf{w}_{d,k}\|^2}{1 - \sigma_e^2 \text{tr}(\mathbf{W}_f^H \mathbf{W}_f)}, \quad (12)$$

which shows that the power of the precoding matrix \mathbf{W}_f for ICI forwarding must be limited by

$$\text{tr}(\mathbf{W}_f^H \mathbf{W}_f) < \frac{1}{\sigma_e^2}. \quad (13)$$

Otherwise, amplifier self-oscillations will occur at the FD PBS [7].

Let P_0 denote the maximal transmit power of the PBS. Then the transmit power P_{out} needs to satisfy $P_{out} \leq P_0$, which can be rewritten based on (12) as

$$\text{tr}(\bar{\mathbf{H}}_{MP} \mathbf{W}_f^H \mathbf{W}_f \bar{\mathbf{H}}_{MP}^H) + (P_0 \sigma_e^2 + \sigma_n^2) \text{tr}(\mathbf{W}_f^H \mathbf{W}_f) + \sum_{k=1}^{K_P} \|\mathbf{w}_{d,k}\|^2 \leq P_0. \quad (14)$$

One can observe that the self-oscillation constraint (13) is implicitly reflected in the maximal power constraint (14). Therefore, in the following only the constraint in (14) is considered.

B. Signals of PUE $_k$

The receive signal of PUE $_k$ can be expressed as

$$y_k[t] = \mathbf{h}_{P_k}^H \mathbf{x}_p[t] + \mathbf{h}_{Mk}^H \mathbf{W}_M \mathbf{s}_M[t] + n_k[t], \quad (15)$$

where the impact of hardware impairments at the HD PUE is ignored as commonly considered for HD transmission in the literature, and $n_k[t] \sim \mathcal{CN}(0, \sigma_n^2)$ is the AWGN at PUE $_k$, which includes the received ICI from interfering PBSs and thermal noises.

By noting the equivalence between time delay and phase shift in narrowband systems, we have $\mathbf{s}_M[t - \tau] = \mathbf{s}_M[t] e^{-j\phi}$, where $\phi = 2\pi f_c \tau$ with f_c denoting the carrier frequency.²

²One can further incorporate the propagation delay difference experienced by the forwarded ICI and the direct ICI into τ , which is not considered here because the PUEs are close to the PBS leading to negligible additional propagation delay.

Then based on (8) and (9), we can rewrite (15) as

$$\begin{aligned}
y_k[t] = & \mathbf{h}_{P_k}^H \mathbf{w}_{d,k} s_{p,k}[t] + \underbrace{\sum_{j=1, j \neq k}^{K_P} \mathbf{h}_{P_k}^H \mathbf{w}_{d,j} s_{p,j}[t]}_{\text{Intra-cell interference}} \\
& + \underbrace{\mathbf{h}_{P_k}^H \mathbf{W}_f (-\mathbf{E}_{PP}^H[t-\tau] \mathbf{x}_p[t-\tau] + \mathbf{H}_{PP}^H \mathbf{z}_x[t-\tau] + \mathbf{n}_p[t-\tau] + \mathbf{z}_y[t-\tau])}_{\text{Forwarded residual self-interference and noises}} \\
& + \underbrace{(\bar{\mathbf{h}}_{Mk}^H + \mathbf{h}_{P_k}^H \mathbf{W}_f \bar{\mathbf{H}}_{MP}^H e^{-j\phi}) \mathbf{s}_M[t] + n_k[t]}_{\text{Cross-tier ICI}}, \quad (16)
\end{aligned}$$

where $\bar{\mathbf{h}}_{Mk} \triangleq \mathbf{W}_M^H \mathbf{h}_{Mk}$ is the equivalent channel from the MBS to PUE_k.

To achieve coherent combination between the ICI forwarded by the PBS and that directly received at the PUEs as shown in (16), the PBS needs to implement sample-by-sample ICI forwarding in time domain to reduce the processing delay as considered in [10, 11] for relay systems, where the time-domain FD amplify-and-forward relay was employed to provide co-phasing combining gain at the destination node. This is different from frequency domain forwarding for FD relay systems, e.g., in [7] by the same authors as [10, 11], where the processing delay will be in symbol-level because a complete symbol needs to be received and demodulated before forwarding in the frequency domain.

Similar to the derivations in Appendix A, we can compute the signal and interference power from (16) by taking the expectation with respect to \mathbf{s}_M , $s_{p,k}$, \mathbf{n}_p , \mathbf{z}_x , \mathbf{z}_y , \mathbf{E}_{PP} , and \mathbf{H}_{PP} , and finally obtain the SINR of PUE_k as

$$\begin{aligned}
\text{SINR}_{k,FD} = & \frac{|\mathbf{h}_{P_k}^H \mathbf{w}_{d,k}|^2 / (\sum_{j \neq k} \|\mathbf{h}_{P_k}^H \mathbf{w}_{d,j}\|^2 + \|\bar{\mathbf{h}}_{Mk}^H + \mathbf{h}_{P_k}^H \mathbf{W}_f \bar{\mathbf{H}}_{MP}^H e^{-j\phi}\|^2)}{+ \|\mathbf{h}_{P_k}^H \mathbf{W}_f\|^2 (P_{out} \sigma_e^2 + \sigma_n^2) + \sigma_n^2}, \quad (17)
\end{aligned}$$

where P_{out} is the transmit power of the PBS, which is a function of the precoders \mathbf{W}_f and $\mathbf{w}_{d,k}$ as given in (12).

III. NARROWBAND SINGLE-USER CASE

In this section, we investigate the single-user case, i.e., $K_M = K_P = 1$, to gain insight into the ICI mitigation mechanism of the fICIC scheme.

In this case, the equivalent channel $\bar{\mathbf{H}}_{MP}$ is a row vector and $\bar{\mathbf{h}}_{Mk}$ is a scalar, which are denoted by $\bar{\mathbf{h}}_{MP}^H \in \mathbb{C}^{1 \times N_r}$ and $\bar{h}_{Mk} \in \mathbb{C}^{1 \times 1}$ for clarity, respectively. Moreover, noting that the intra-cell interference does not exist now, the SINR can be simplified as

$$\text{SINR}_{k,FD} = \quad (18)$$

$$\frac{|\mathbf{h}_{P_k}^H \mathbf{w}_{d,k}|^2}{|\bar{h}_{Mk}^* + \mathbf{h}_{P_k}^H \mathbf{W}_f \bar{\mathbf{H}}_{MP}^H e^{-j\phi}|^2 + \|\mathbf{h}_{P_k}^H \mathbf{W}_f\|^2 (P_{out} \sigma_e^2 + \sigma_n^2) + \sigma_n^2},$$

where P_{out} given in (12) can be rewritten as

$$P_{out} = \frac{\|\mathbf{W}_f \bar{\mathbf{H}}_{MP}\|^2 + \sigma_n^2 \text{tr}(\mathbf{W}_f^H \mathbf{W}_f) + \|\mathbf{w}_{d,k}\|^2}{1 - \sigma_e^2 \text{tr}(\mathbf{W}_f^H \mathbf{W}_f)}. \quad (19)$$

When the precoder for forwarding the listened ICI, \mathbf{W}_f , in (18) is selected as zero, the FD PBS reduces to a HD PBS. It is not hard to obtain the optimal precoder for transmitting the desired signal in HD case that maximizes the SINR of PUE_k as $\mathbf{w}_{d,k} = \frac{\sqrt{P_0} \mathbf{h}_{P_k}}{\|\mathbf{h}_{P_k}\|}$, which is referred to as “the HD scheme” in the sequel. The maximum SINR achieved by the HD scheme can be obtained from (18) as

$$\text{SINR}_{k,HD}^* = \frac{P_0 \|\mathbf{h}_{P_k}\|^2}{|\bar{h}_{Mk}|^2 + \sigma_n^2}. \quad (20)$$

Compared to the HD scheme where the ICI power is $|\bar{h}_{Mk}|^2$ as shown in (20), the FD scheme turns the ICI power into $|\bar{h}_{Mk}^* + \mathbf{h}_{P_k}^H \mathbf{W}_f \bar{\mathbf{H}}_{MP}^H e^{-j\phi}|^2$ as shown in (18), which can be reduced by optimizing the precoders \mathbf{W}_f and $\mathbf{w}_{d,k}$. The optimization problem, aimed at maximizing the SINR of PUE_k subject to the maximal transmit power constraint, can be formulated as

$$\max_{\mathbf{W}_f, \mathbf{w}_{d,k}} \text{SINR}_{k,FD} \quad (21a)$$

$$\begin{aligned}
s. \quad & \|\mathbf{W}_f \bar{\mathbf{H}}_{MP}\|^2 + (P_0 \sigma_e^2 + \sigma_n^2) \text{tr}(\mathbf{W}_f^H \mathbf{W}_f) \\
& + \|\mathbf{w}_{d,k}\|^2 \leq P_0, \quad (21b)
\end{aligned}$$

where the power constraint (21b) comes from (14).

A. Optimal fICIC Scheme

In this subsection, we strive to find the optimal fICIC scheme with explicit expressions for \mathbf{W}_f and $\mathbf{w}_{d,k}$. Given that it is difficult to directly solve problem (21) due to the complex expression of $\text{SINR}_{k,FD}$ in (18), we start with investigating the properties of the optimal solutions to problem (21) in the following.

First, by substituting the expression of P_{out} given in (19) into (18), we can find that $\text{SINR}_{k,FD}$, i.e., the objective function of problem (21), is an increasing function of $\|\mathbf{w}_{d,k}\|^2$. Therefore, for any given \mathbf{W}_f and the direction of $\mathbf{w}_{d,k}$, $\frac{\mathbf{w}_{d,k}}{\|\mathbf{w}_{d,k}\|}$, we can always improve the SINR by increasing $\|\mathbf{w}_{d,k}\|^2$ until the PBS transmits with its maximum power. This means that the global optimal solution to problem (21) is obtained when the power constraint in (21b) holds with equality.

Based on this result, we can obtain from (19) that $P_{out} = P_0$, with which $\text{SINR}_{k,FD}$ is simplified as

$$\text{SINR}_{k,FD} = \quad (22)$$

$$\frac{|\mathbf{h}_{P_k}^H \mathbf{w}_{d,k}|^2}{|\bar{h}_{Mk}^* + \mathbf{h}_{P_k}^H \mathbf{W}_f \bar{\mathbf{H}}_{MP}^H e^{-j\phi}|^2 + \|\mathbf{h}_{P_k}^H \mathbf{W}_f\|^2 (P_0 \sigma_e^2 + \sigma_n^2) + \sigma_n^2}.$$

Second, because the direction of $\mathbf{w}_{d,k}$ affects only the numerator of $\text{SINR}_{k,FD}$, we can obtain that the optimal $\mathbf{w}_{d,k}^*$ satisfies

$$\frac{\mathbf{w}_{d,k}^*}{\|\mathbf{w}_{d,k}^*\|} = \frac{\mathbf{h}_{P_k}}{\|\mathbf{h}_{P_k}\|}. \quad (23)$$

Third, we can show that the optimal \mathbf{W}_f^* has the following property.

Lemma 1: The optimal \mathbf{W}_f^* is of rank 1, which can be decomposed as

$$\mathbf{W}_f^* = -\bar{h}_{Mk}^* e^{j\phi} \cdot \beta \cdot \mathbf{h}_{P_k} \bar{\mathbf{H}}_{MP}^H, \quad (24)$$

where β is a positive scalar.

Proof: See Appendix B. ■

We can see from the expression of the optimal \mathbf{W}_f^* that with the fICIC the listened ICI is first enhanced by the maximal ratio combining with $\bar{\mathbf{h}}_{MP}^H$, and then forwarded by the maximal ratio transmission with \mathbf{h}_{Pk} .

Based on (23) and (24), we can express the SINR with β and $\|\mathbf{w}_{d,k}\|$. Further considering that the optimal solution is obtained when the power constraint in (21b) holds with equality, we can replace $\|\mathbf{w}_{d,k}\|$ with β , and finally convert problem (21) into the following optimization problem

$$\max_{\beta} f_0(\beta) \quad (25a)$$

$$s. t. \quad (\|\bar{\mathbf{h}}_{MP}\|^2 + \sigma_I^2 + \sigma_n^2) |\bar{h}_{Mk}|^2 \|\mathbf{h}_{Pk}\|^2 \|\bar{\mathbf{h}}_{MP}\|^2 \beta^2 \leq P_0 \quad (25b)$$

$$\beta \geq 0, \quad (25c)$$

where $f_0(\beta) \triangleq \left(\|\mathbf{h}_{Pk}^H\|^2 (P_0 - (\|\bar{\mathbf{h}}_{MP}\|^2 + \sigma_I^2 + \sigma_n^2) |\bar{h}_{Mk}|^2 \|\mathbf{h}_{Pk}\|^2 \|\bar{\mathbf{h}}_{MP}\|^2 \beta^2) \right) / \left((|\bar{h}_{Mk}| - \beta |\bar{h}_{Mk}| \|\mathbf{h}_{Pk}\|^2 \|\bar{\mathbf{h}}_{MP}\|^2)^2 + \beta^2 |\bar{h}_{Mk}|^2 \|\mathbf{h}_{Pk}\|^4 \|\bar{\mathbf{h}}_{MP}\|^2 (\sigma_I^2 + \sigma_n^2) + \sigma_n^2 \right)$, and $\sigma_I^2 \triangleq P_0 \sigma_e^2$ denotes the average power of residual self-interference.

Proposition 1: The maximum of the objective function in (25a), i.e. the maximal SINR of PUE_k, can be expressed as

$$\text{SINR}_{k,FD}^* = \frac{1}{\frac{1}{B\beta^*} - 1}, \quad (26)$$

with $\beta^* = \frac{2(A+D-\sqrt{(A+D)^2 - \frac{AC^2}{B}})}{C^2}$ representing the optimal solution of β , where $A = P_0 \|\mathbf{h}_{Pk}\|^2$, $B = \|\mathbf{h}_{Pk}\|^2 (\|\bar{\mathbf{h}}_{MP}\|^2 + \sigma_I^2 + \sigma_n^2)$, $C = 2|\bar{h}_{Mk}| \|\mathbf{h}_{Pk}\| \|\bar{\mathbf{h}}_{MP}\|$, and $D = |\bar{h}_{Mk}|^2 + \sigma_n^2$.

Proof: See Appendix C. ■

With the optimal β^* , we can directly obtain the optimal \mathbf{W}_f^* from (24). To compute the optimal $\mathbf{w}_{d,k}^*$, recalling that constraint (21b) holds with equality for the optimal solutions, we can obtain the norm of the optimal $\mathbf{w}_{d,k}^*$ as

$$\|\mathbf{w}_{d,k}^*\| = \sqrt{P_0 - \frac{1}{4} (\|\bar{\mathbf{h}}_{MP}\|^2 + \sigma_I^2 + \sigma_n^2) C^2 \beta^{*2}}. \quad (27)$$

Then by substituting (27) into (23), we can obtain $\mathbf{w}_{d,k}^*$.

B. Asymptotic Performance Analysis

To understand the behavior of the fICIC scheme, we next consider an ICI-dominated scenario, where the noise power σ_n^2 approaches to zero.

1) Perfect Self-interference Cancellation

If the self-interference can be perfectly eliminated, i.e., $\sigma_I^2 = 0$, from the definitions after (26) we have $B \doteq \|\mathbf{h}_{Pk}\|^2 \|\bar{\mathbf{h}}_{MP}\|^2$ and $D \doteq |\bar{h}_{Mk}|^2$, where \doteq denotes asymptotic equality. Then based on Proposition 1, after some manipulations we can obtain the optimal β^* as

$$\begin{aligned} \beta^* &\doteq \frac{P_0 \|\mathbf{h}_{Pk}\|^2 + |\bar{h}_{Mk}|^2 - |P_0 \|\mathbf{h}_{Pk}\|^2 - |\bar{h}_{Mk}|^2|}{2 \|\bar{\mathbf{h}}_{MP}\|^2 |\bar{h}_{Mk}|^2 \|\mathbf{h}_{Pk}\|^2} \quad (28) \\ &= \frac{\min(P_0 \|\mathbf{h}_{Pk}\|^2, |\bar{h}_{Mk}|^2)}{\|\bar{\mathbf{h}}_{MP}\|^2 |\bar{h}_{Mk}|^2 \|\mathbf{h}_{Pk}\|^2} \triangleq \frac{\eta}{\|\bar{\mathbf{h}}_{MP}\|^2 |\bar{h}_{Mk}|^2 \|\mathbf{h}_{Pk}\|^2}. \end{aligned}$$

By substituting (28) into (25a), the maximal SINR of PUE_k becomes

$$\text{SINR}_{k,FD}^* \doteq \frac{P_0 |\bar{h}_{Mk}|^2 \|\mathbf{h}_{Pk}\|^2 - \eta^2}{(\eta - |\bar{h}_{Mk}|^2)^2 + \left(\frac{\eta^2}{\|\bar{\mathbf{h}}_{MP}\|^2} + |\bar{h}_{Mk}|^2 \right) \sigma_n^2}, \quad (29)$$

where $\eta = \min(P_0 \|\mathbf{h}_{Pk}\|^2, |\bar{h}_{Mk}|^2)$ as defined in (28), which depends on the strengths of the desired signal, $P_0 \|\mathbf{h}_{Pk}\|^2$, and the ICI, $|\bar{h}_{Mk}|^2$. In the following two cases are discussed.

• Case 1: $|\bar{h}_{Mk}|^2 < P_0 \|\mathbf{h}_{Pk}\|^2$

This is a case where the ICI is weaker than the desired signal. Then, we have $\eta = |\bar{h}_{Mk}|^2$ and the maximal SINR is

$$\overline{\text{SINR}}_{k,FD}^* \doteq \frac{P_0 \|\mathbf{h}_{Pk}\|^2 - |\bar{h}_{Mk}|^2}{\left(\frac{|\bar{h}_{Mk}|^2}{\|\bar{\mathbf{h}}_{MP}\|^2} + 1 \right) \sigma_n^2}. \quad (30)$$

It implies that the FD PBS can thoroughly eliminate the ICI generated by the MBS by properly designing the forwarding and transmitting precoders.

When compared with the HD scheme, the performance gain of the fICIC can be obtained as

$$\frac{\overline{\text{SINR}}_{k,FD}^*}{\text{SINR}_{k,HD}^*} \doteq \frac{1 - \frac{|\bar{h}_{Mk}|^2}{P_0 \|\mathbf{h}_{Pk}\|^2}}{\left(\frac{1}{\|\bar{\mathbf{h}}_{MP}\|^2} + \frac{1}{|\bar{h}_{Mk}|^2} \right) \sigma_n^2}. \quad (31)$$

• Case 2: $|\bar{h}_{Mk}|^2 \geq P_0 \|\mathbf{h}_{Pk}\|^2$

In this case where the ICI is stronger, we have $\eta = P_0 \|\mathbf{h}_{Pk}\|^2$ and the maximal SINR is

$$\widehat{\text{SINR}}_{k,FD}^* \doteq \frac{P_0 \|\mathbf{h}_{Pk}\|^2}{|\bar{h}_{Mk}|^2 - P_0 \|\mathbf{h}_{Pk}\|^2 + \frac{P_0^2 \|\mathbf{h}_{Pk}\|^4 + |\bar{h}_{Mk}|^2}{\|\bar{\mathbf{h}}_{MP}\|^2 + |\bar{h}_{Mk}|^2} \sigma_n^2}. \quad (32)$$

For very strong interference, i.e., $|\bar{h}_{Mk}|^2 \gg P_0 \|\mathbf{h}_{Pk}\|^2$, $\widehat{\text{SINR}}_{k,FD}^*$ can be approximated as

$$\widehat{\text{SINR}}_{k,FD}^* \approx \frac{P_0 \|\mathbf{h}_{Pk}\|^2}{|\bar{h}_{Mk}|^2} \doteq \text{SINR}_{k,HD}^*. \quad (33)$$

From the above analysis, we can obtain the following observations.

• **Impact of $\bar{\mathbf{h}}_{MP}$:** It is shown from (30) and (32) that the SINR of PUE_k increases with the channel gain between MBS and PBS $\|\bar{\mathbf{h}}_{MP}\|$. This is because given the power of the PBS allocated for forwarding the listened ICI, denoted as

$$P_{out}^{fw} \triangleq \|\mathbf{W}_f \bar{\mathbf{h}}_{MP}\|^2 + (\sigma_I^2 + \sigma_n^2) \text{tr}(\mathbf{W}_f^H \mathbf{W}_f), \quad (34)$$

a large $\|\bar{\mathbf{h}}_{MP}\|$ will reduce the power used for forwarding the residual self-interference and noises and thus improve the efficiency of power usage. However, when the ICI power $|\bar{h}_{Mk}|^2$ is very large, the impact of $\|\bar{\mathbf{h}}_{MP}\|$ can be neglected as shown in (33).

• **Impact of \mathbf{h}_{Pk} and \bar{h}_{Mk} :** As shown by (20), (30) and (33), increasing the desired signal $P_0 \|\mathbf{h}_{Pk}\|^2$ can improve the performance of both the FD scheme and the HD

scheme.³ However, the performance improvement for the FD scheme is more significant because (31) shows that the performance gain of FD over HD increases with $P_0 \|\mathbf{h}_{Pk}\|^2$. It can also be seen that the performance gain of FD over HD decreases with the ICI power $|\bar{h}_{Mk}|^2$, until vanishes for very large $|\bar{h}_{Mk}|^2$ as shown by (33). It should be pointed out that the extreme case with very strong ICI in (33) rarely happens in practice even when the maximum CRE offset of 9 dB in LTE systems is considered [12]. As will be verified in Fig. 4, the fICIC still exhibits evident performance gain over the HD scheme when the average power of the ICI is 10.4 dB stronger than the desired signal.

2) Large Residual Self-interference

When the self-interference cancellation for FD is imperfect and the residual self-interference, σ_I^2 , is large, the parameter B is large and the term $\frac{AC^2}{B}$ is small. Then by using the first-order Taylor expansion, $\sqrt{c-z} \doteq \sqrt{c} - \frac{1}{2\sqrt{c}}z$ for small z , we can obtain $\beta^* \doteq \frac{A}{(A+D)B}$ when $\frac{AC^2}{B}$ approaches to zero, with which we can obtain from (26) the maximal SINR of PUE_k as

$$\widetilde{\text{SINR}}_{k,FD}^* \doteq \frac{P_0 \|\mathbf{h}_{Pk}\|^2}{|\bar{h}_{Mk}|^2 + \sigma_n^2} = \text{SINR}_{k,HD}^*. \quad (35)$$

Moreover, we can compute the power of the PBS allocated for forwarding the listened ICI from (34) as

$$P_{out}^{fw} \doteq \frac{P_0^2 \|\bar{\mathbf{h}}_{MP}\|^2 |\bar{h}_{Mk}|^2 \|\mathbf{h}_{Pk}\|^2}{(|\bar{h}_{Mk}|^2 + P_0 \|\mathbf{h}_{Pk}\|^2 + \sigma_n^2)^2 (\|\bar{\mathbf{h}}_{MP}\|^2 + \sigma_I^2 + \sigma_n^2)}. \quad (36)$$

It can be seen from (36) that the transmit power of the PBS allocated for forwarding ICI decreases with the growth of residual self-interference σ_I^2 . When σ_I^2 is very large, the PBS will use all power to transmit desired signals, and therefore the fICIC will reduce to the HD scheme.

IV. NARROWBAND MULTI-USER CASE

In this section, we consider the general multi-user case, where the MBS serves K_M MUEs and the PBS serves K_P PUEs with $K_M \geq 1$ and $K_P \geq 1$.

We optimize the fICIC scheme, aimed at maximizing the sum rate of K_P PUEs served by the reference PBS while guaranteeing the fairness among the PUEs, subject to the maximal transmit power constraint of the PBS. The problem can be formulated as

$$\max_{\mathbf{w}_f, \{\mathbf{w}_{d,k}\}} R_{sum} \quad (37a)$$

$$s. t. \log(1 + \text{SINR}_{k,FD}) = \alpha_k R_{sum}, \quad k = 1, \dots, K_P \quad (37b)$$

$$\begin{aligned} & \text{tr}(\bar{\mathbf{H}}_{MP} \mathbf{W}_f^H \mathbf{W}_f \bar{\mathbf{H}}_{MP}^H) + (\sigma_I^2 + \sigma_n^2) \text{tr}(\mathbf{W}_f^H \mathbf{W}_f) \\ & + \sum_{k=1}^{K_P} \|\mathbf{w}_{d,k}\|^2 \leq P_0, \quad (37c) \end{aligned}$$

³Herein, we consider that the increase of the desired signal's strength comes from reducing the distance between the PUE and its serving PBS, but not from increasing the power of the PBS. As a result, the interference from surrounding PBSs to the PUE will be even weaker, making it reasonable to focus on the dominant cross-tier ICI as we considered.

where R_{sum} is the sum rate of all PUEs, the constraints in (37b) ensure the data rate proportion among the PUEs with predefined fairness factors $\{\alpha_k\}$ as considered in [13], $\alpha_k \geq 0$, and $\sum_{k=1}^{K_P} \alpha_k = 1$.

The expression of $\text{SINR}_{k,FD}$ in multi-user case is given in (17), which is a function of the transmit power P_{out} as shown in (12) and thus is very complicated. To simplify $\text{SINR}_{k,FD}$, we can prove that the optimal solution to problem (37) is obtained when the power constraint (37c) holds with equality.⁴ Based on this result, we have $P_{out} = P_0$ in (17).

Problem (37) is to find the maximal achievable R_{sum} , denoted by R_{sum}^* , ensuring all constraints satisfied, which can be solved by bisection methods [14]. Specifically, for a given R_{sum} in an iteration, denoted by R_{sum}^0 , if the following optimization problem

$$\text{Find } \mathbf{W}_f, \{\mathbf{w}_{d,k}\} \quad (38a)$$

$$s. t. \quad (37c)$$

$$\text{SINR}_{k,FD} = 2^{\alpha_k R_{sum}^0} - 1, \quad k = 1, \dots, K_P \quad (38b)$$

is feasible, then it follows that R_{sum}^0 is an achievable sum rate of all K_P PUEs, i.e., $R_{sum}^0 \leq R_{sum}^*$, otherwise, $R_{sum}^0 > R_{sum}^*$. This condition can be used in bisection algorithms to find R_{sum}^* .

Now the remaining issue is to find efficient approaches to evaluate the feasibility of problem (38). In the following, we show that the feasibility problem can be solved by investigating the optimization problem below

$$\begin{aligned} & \min_{\mathbf{w}_f, \{\mathbf{w}_{d,k}\}} \text{tr}(\bar{\mathbf{H}}_{MP} \mathbf{W}_f^H \mathbf{W}_f \bar{\mathbf{H}}_{MP}^H) + (\sigma_I^2 + \sigma_n^2) \cdot \\ & \text{tr}(\mathbf{W}_f^H \mathbf{W}_f) + \sum_{k=1}^{K_P} \|\mathbf{w}_{d,k}\|^2 \quad (39a) \end{aligned}$$

$$s. t. \quad \text{SINR}_{k,FD} \geq \gamma_k, \quad k = 1, \dots, K_P, \quad (39b)$$

where the objective function is the left-hand side of constraint (37c), and $\gamma_k \triangleq 2^{\alpha_k R_{sum}^0} - 1$.

To see this, we note that the optimal solution to problem (39) is obtained when all SINR constraints in (39b) hold with equality, otherwise, the value of the objective function can be always further reduced by properly decreasing $\|\mathbf{w}_{d,k}\|^2$. It suggests that if the minimum of the objective function (39a) is smaller than P_0 , then constraints (37c) and (38b) in problem (38) are all satisfied. This means that problem (38) is feasible and R_{sum}^0 is an achievable sum rate, otherwise, problem (38) will be infeasible. In what follows, we solve problem (39).

By defining $\mathbf{w}_f = \text{vec}(\mathbf{W}_f^H)$, we can rewrite problem (39) as

$$\begin{aligned} & \min_{\mathbf{w}_f, \{\mathbf{w}_{d,k}\}} \|(\mathbf{I}_{N_t} \otimes \bar{\mathbf{H}}_{MP}) \mathbf{w}_f\|^2 + (\sigma_I^2 + \sigma_n^2) \|\mathbf{w}_f\|^2 + \sum_{k=1}^{K_P} \|\mathbf{w}_{d,k}\|^2 \\ & \quad (40a) \end{aligned}$$

$$s. t. \quad \frac{|\mathbf{h}_{Pk}^H \mathbf{w}_{d,k}|^2}{\Omega_k} \geq \gamma_k, \quad \forall k, \quad (40b)$$

⁴Otherwise, suppose that the power constraint holds with inequality with the optimal precoders \mathbf{W}_f^* and $\{\mathbf{w}_{d,k}^*\}$, then given \mathbf{W}_f^* , we can always find new precoders $\{\mathbf{w}_{d,k}^{*'}\}$, defined as $\mathbf{w}_{d,k}^{*'} = c \mathbf{w}_{d,k}^*$ for $k = 1, \dots, K_P$ with $c > 1$, which can further improve the SINR of all PUEs until the constraint (37c) holds with equality.

where $\Omega_k \triangleq \sum_{j \neq k} \|\mathbf{h}_{P_k}^H \mathbf{w}_{d,j}\|^2 + \|\bar{\mathbf{h}}_{Mk} + e^{-j\phi} (\mathbf{h}_{P_k}^T \otimes \bar{\mathbf{H}}_{MP}) \mathbf{w}_f\|^2 + \|(\mathbf{h}_{P_k}^T \otimes \mathbf{I}_{N_r}) \mathbf{w}_f\|^2 (\sigma_I^2 + \sigma_n^2) + \sigma_n^2$.

Problem (40) is non-convex because the constraints in (40b) are non-convex. A common method to solve such a problem is to convert the non-convex SINR constraints into convex second-order cone constraints [15]. Specifically, since adding any phase rotation to $\mathbf{w}_{d,k}$ will not affect the SINR of all PUEs, we can assume that $\mathbf{h}_{P_k}^H \mathbf{w}_{d,k}$ is real-valued, which does not affect the global optima of problem (40). Then we can convert the non-convex problem (40) into the following second-order cone constrained problem

$$\min_{\mathbf{w}_f, \mathbf{w}_d} \|(\mathbf{I}_{N_t} \otimes \bar{\mathbf{H}}_{MP}) \mathbf{w}_f\|^2 + (\sigma_I^2 + \sigma_n^2) \|\mathbf{w}_f\|^2 + \|\mathbf{w}_d\|^2 \quad (41a)$$

$$\begin{aligned} s. t. \quad & \sqrt{1 + \frac{1}{\gamma_k} \mathbf{h}_{P_k}^H \mathbf{w}_{d,k}} \\ & \geq \left\| \begin{bmatrix} (\mathbf{I}_{N_r} \otimes \mathbf{h}_{P_k}^H) \mathbf{w}_d \\ e^{-j\phi} (\mathbf{h}_{P_k}^T \otimes \bar{\mathbf{H}}_{MP}) \mathbf{w}_f \\ \sqrt{\sigma_I^2 + \sigma_n^2} (\mathbf{h}_{P_k}^T \otimes \mathbf{I}_{N_r}) \mathbf{w}_f \end{bmatrix} + \begin{bmatrix} \bar{\mathbf{0}}_{N_r} \\ \bar{\mathbf{h}}_{Mk} \\ \bar{\mathbf{0}}_{N_r} \\ 0 \end{bmatrix} \right\|, \forall k, \end{aligned} \quad (41b)$$

where $\mathbf{w}_d = \text{vec}([\mathbf{w}_{d,1}, \dots, \mathbf{w}_{d,K_P}])$.

The resultant problem (41) is convex since both the objective function and constraints are convex, which can be solved by standard convex optimization algorithms [14]. However, the computational complexity of the standard algorithms is still too high and prohibits the practical use of the fICIC scheme, especially when the numbers of MUEs, PUEs, and the transmit and receive antennas at the PBS are large. In the following we strive to propose a low-complexity algorithm to solve problem (40), with which the optimal precoders are obtained with explicit expressions. We begin with the discussion regarding the strong duality of problem (40).

Recalling that we have shown the equivalence between problem (40) and problem (41), and also known that strong duality holds for problem (41) because it is a convex problem. Further, along the lines of [15, App.A], we can show the equivalence between the Lagrangian functions of problem (41) and problem (40). Therefore, strong duality holds for the non-convex optimization problem (40). This means that we can solve problem (40) with the Lagrange dual method.

The dual problem to problem (40) can be expressed as

$$\max_{\bar{\lambda}_k} \min_{\mathbf{w}_f, \{\mathbf{w}_{d,k}\}} J(\bar{\lambda}_k, \mathbf{w}_f, \mathbf{w}_{d,k}) \quad (42a)$$

$$s. t. \quad \bar{\lambda}_k \geq 0, \forall k, \quad (42b)$$

where $\bar{\lambda}_k$ is the lagrangian multiplier, and $J(\bar{\lambda}_k, \mathbf{w}_f, \mathbf{w}_{d,k})$ is the Lagrangian function of problem (40) with the expression

$$\begin{aligned} J(\bar{\lambda}_k, \mathbf{w}_f, \mathbf{w}_{d,k}) = & \|(\mathbf{I}_{N_t} \otimes \bar{\mathbf{H}}_{MP}) \mathbf{w}_f\|^2 + (\sigma_I^2 + \sigma_n^2) \|\mathbf{w}_f\|^2 \\ & + \sum_{k=1}^{K_P} \bar{\lambda}_k (\|\bar{\mathbf{h}}_{Mk} + e^{-j\phi} (\mathbf{h}_{P_k}^T \otimes \bar{\mathbf{H}}_{MP}) \mathbf{w}_f\|^2 \\ & + \|(\mathbf{h}_{P_k}^T \otimes \mathbf{I}_{N_r}) \mathbf{w}_f\|^2 (\sigma_I^2 + \sigma_n^2) + \sigma_n^2) \\ & + \sum_{k=1}^{K_P} \mathbf{w}_{d,k}^H \left(\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j \mathbf{h}_{P_j} \mathbf{h}_{P_j}^H - \frac{\bar{\lambda}_k}{\gamma_k} \mathbf{h}_{P_k} \mathbf{h}_{P_k}^H \right) \mathbf{w}_{d,k}. \end{aligned} \quad (43)$$

A. Optimal Solution to $\bar{\lambda}_k$

It can be seen from (43) that $\min_{\mathbf{w}_f, \{\mathbf{w}_{d,k}\}} J(\bar{\lambda}_k, \mathbf{w}_f, \mathbf{w}_{d,k}) \rightarrow -\infty$ except when $\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j \mathbf{h}_{P_j} \mathbf{h}_{P_j}^H - \frac{\bar{\lambda}_k}{\gamma_0} \mathbf{h}_{P_k} \mathbf{h}_{P_k}^H \succeq \mathbf{0}_{N_t}$. Hence, the dual problem (42) is equivalent to

$$\min_{\mathbf{w}_f} \max_{\bar{\lambda}_k} J(\bar{\lambda}_k, \mathbf{w}_f, \bar{\mathbf{0}}_{N_t}) \quad (44a)$$

$$s. t. \quad \mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j \mathbf{h}_{P_j} \mathbf{h}_{P_j}^H - \frac{\bar{\lambda}_k}{\gamma_k} \mathbf{h}_{P_k} \mathbf{h}_{P_k}^H \succeq \mathbf{0}_{N_t}, \forall k \quad (44b)$$

$$\bar{\lambda}_k \geq 0, \forall k, \quad (44c)$$

where the objective function (44a) comes from the fact that $J(\bar{\lambda}_k, \mathbf{w}_f, \mathbf{w}_{d,k})$ given in (43) is minimized when $\mathbf{w}_{d,k} = \bar{\mathbf{0}}_{N_t}$.

It is difficult to directly find the optimal $\bar{\lambda}_k$ from problem (44) due to the complicated semi-definite positive constraints (44b). To solve the problem, we simplify the constraints as follows.

Proposition 2: The semi-definite positive constraints (44b) can be equivalently expressed as

$$\bar{\lambda}_k \mathbf{h}_{P_k}^H \left(\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j \mathbf{h}_{P_j} \mathbf{h}_{P_j}^H \right)^{-1} \mathbf{h}_{P_k} \leq \gamma_k, \forall k. \quad (45)$$

Proof: See Appendix D. ■

Note that for any given \mathbf{w}_f , the objective function of problem (44) is an increasing function of $\bar{\lambda}_k$. Then we can show that the optimal $\bar{\lambda}_k$ maximizing $J(\bar{\lambda}_k, \mathbf{w}_f, \bar{\mathbf{0}}_{N_t})$ is obtained when the constraints in (45) hold with equality (otherwise, one can always increase $\bar{\lambda}_k$ to improve the value of the objective function). It suggests that the optimal $\bar{\lambda}_k^*$ should satisfy

$$\bar{\lambda}_k^* = \frac{\gamma_k}{\mathbf{h}_{P_k}^H (\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j^* \mathbf{h}_{P_j} \mathbf{h}_{P_j}^H)^{-1} \mathbf{h}_{P_k}}. \quad (46)$$

(46) provides a fixed-point iterative algorithm to find the optimal $\bar{\lambda}_k^*$, which can be expressed as

$$\bar{\lambda}_k^{*(n+1)} = \frac{\gamma_k}{\mathbf{h}_{P_k}^H (\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j^{*(n)} \mathbf{h}_{P_j} \mathbf{h}_{P_j}^H)^{-1} \mathbf{h}_{P_k}}, \quad (47)$$

where the superscript (n) denotes the n -th iteration.

The convergence of the fixed-point iterative algorithm can be proved based on the standard function theory [16], which shows that the algorithm given in (47) will converge to a unique optimal solution from any initial value $\{\bar{\lambda}_k^{*(0)}\}$.

B. Optimal Solution to \mathbf{w}_f

After obtaining the optimal $\bar{\lambda}_k^*$ for $k = 1, \dots, K_P$, we next find the optimal \mathbf{w}_f^* based on the Karush-Kuhn-Tucker (KKT) condition [14] associated with the Lagrangian function $J(\bar{\lambda}_k^*, \mathbf{w}_f, \mathbf{w}_{d,k})$. We can obtain that the optimal \mathbf{w}_f^* satisfies

$$\begin{aligned} & \sum_{k=1}^{K_P} e^{j\phi} \bar{\lambda}_k^* (\mathbf{h}_{P_k}^T \otimes \bar{\mathbf{H}}_{MP})^H \bar{\mathbf{h}}_{Mk} + (\mathbf{I}_{N_t} \otimes \bar{\mathbf{H}}_{MP}^H \bar{\mathbf{H}}_{MP}) \mathbf{w}_f^* \\ & + (\sigma_I^2 + \sigma_n^2) \mathbf{w}_f^* + \sum_{k=1}^{K_P} \bar{\lambda}_k^* ((\mathbf{h}_{P_k} \mathbf{h}_{P_k}^H)^T \otimes \bar{\mathbf{H}}_{MP}^H \bar{\mathbf{H}}_{MP}) \mathbf{w}_f^* \\ & + (\sigma_I^2 + \sigma_n^2) ((\mathbf{h}_{P_k} \mathbf{h}_{P_k}^H)^T \otimes \mathbf{I}_{N_r}) \mathbf{w}_f^* = \bar{\mathbf{0}}_{N_t N_r}. \end{aligned} \quad (48)$$

By applying the properties of Kronecker product [17] and after some manipulations, we can obtain the optimal \mathbf{w}_f^* as

$$\mathbf{w}_f^* = -e^{j\phi} \left(\left(\sum_{k=1}^{K_P} \bar{\lambda}_k^* (\mathbf{h}_{P_k} \mathbf{h}_{P_k}^H)^T + \mathbf{I}_{N_t} \right)^{-1} \left(\sum_{k=1}^{K_P} \bar{\lambda}_k^* \bar{\mathbf{h}}_{Mk} \mathbf{h}_{P_k}^H \right)^T \right. \\ \left. \otimes (\bar{\mathbf{H}}_{MP}^H \bar{\mathbf{H}}_{MP} + (\sigma_I^2 + \sigma_n^2) \mathbf{I}_{N_r})^{-1} \bar{\mathbf{H}}_{MP}^H \right) \text{vec}(\mathbf{I}_{N_t N_r}). \quad (49)$$

Then we can recover the optimal \mathbf{W}_f^* from \mathbf{w}_f^* as

$$\mathbf{W}_f^* = -e^{j\phi} \left(\sum_{k=1}^{K_P} \bar{\lambda}_k^* \mathbf{h}_{P_k} \mathbf{h}_{P_k}^H + \mathbf{I}_{N_t} \right)^{-1}. \quad (50)$$

$$\left(\sum_{k=1}^{K_P} \bar{\lambda}_k^* \mathbf{h}_{P_k} \bar{\mathbf{h}}_{Mk}^H \right) \bar{\mathbf{H}}_{MP} (\bar{\mathbf{H}}_{MP}^H \bar{\mathbf{H}}_{MP} + (\sigma_I^2 + \sigma_n^2) \mathbf{I}_{N_r})^{-1}.$$

C. Optimal Solution to $\mathbf{w}_{d,k}$

With the optimal λ_k^* and \mathbf{W}_f^* , the optimal $\mathbf{w}_{d,k}^*$ of problem (40) can be solved as follows.

According to the KKT condition, the optimal $\mathbf{w}_{d,k}^*$ satisfies

$$\frac{\partial J(\bar{\lambda}_k^*, \mathbf{w}_f^*, \mathbf{w}_{d,k})}{\partial \mathbf{w}_{d,k}} \\ = 2 \left(\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j^* \mathbf{h}_{P,j} \mathbf{h}_{P,j}^H - \frac{\bar{\lambda}_k^*}{\gamma_k} \mathbf{h}_{P_k} \mathbf{h}_{P_k}^H \right) \mathbf{w}_{d,k} = \bar{\mathbf{0}}_{N_t}, \quad (51)$$

from which we have

$$\mathbf{w}_{d,k}^* = \frac{\lambda_k^* \mathbf{h}_{P_k}^H \mathbf{w}_{d,k}^*}{\gamma_k} (\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j^* \mathbf{h}_{P,j} \mathbf{h}_{P,j}^H)^{-1} \mathbf{h}_{P_k} \\ \triangleq \sqrt{p_k^*} \tilde{\mathbf{w}}_{d,k}^*, \quad (52)$$

where $\tilde{\mathbf{w}}_{d,k}^* = (\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j^* \mathbf{h}_{P,j} \mathbf{h}_{P,j}^H)^{-1} \mathbf{h}_{P_k}$, and p_k^* is a scalar controlling the power allocated to PUE_k for transmitting desired signals.

To find the optimal $\{p_k^*\}$ in (52), recalling that the optimal solution to problem (40) is obtained when all SINR constraints in (40b) hold with equality if problem (40) is feasible for given $\{\gamma_k\}$, then we can obtain the following equations with respect to p_k^*

$$\frac{p_k^* |\mathbf{h}_{P_k}^H \tilde{\mathbf{w}}_{d,k}^*|^2}{\sum_{j \neq k} p_j^* \|\mathbf{h}_{P_k}^H \tilde{\mathbf{w}}_{d,j}^*\|^2 + \zeta_k(\mathbf{W}_f^*)} = \gamma_k, \quad k=1, \dots, K_P, \quad (53)$$

where $\zeta_k(\mathbf{W}_f^*) = \|\bar{\mathbf{h}}_{Mk}^H + \mathbf{h}_{P_k}^H \mathbf{W}_f^* \bar{\mathbf{H}}_{MP}^H e^{-j\phi}\|^2 + \|\mathbf{h}_{P_k}^H \mathbf{W}_f^*\|^2 (\sigma_I^2 + \sigma_n^2) + \sigma_n^2$.

From the equations in (53), we can solve the optimal $\mathbf{p}^* \triangleq [p_1^*, \dots, p_{K_P}^*]^T$ as

$$\mathbf{p}^* = \mathbf{M}^{-1} \boldsymbol{\zeta}, \quad (54)$$

where $\boldsymbol{\zeta} = [\zeta_1(\mathbf{W}_f^*), \dots, \zeta_{K_P}(\mathbf{W}_f^*)]^T$, and $\mathbf{M} \in \mathbb{C}^{K_P \times K_P}$ is defined as

$$[\mathbf{M}]_{kj} = \begin{cases} \frac{1}{\gamma_k} |\mathbf{h}_{P_k}^H \tilde{\mathbf{w}}_{d,k}^*|^2, & k=j, \\ -|\mathbf{h}_{P_k}^H \tilde{\mathbf{w}}_{d,j}^*|^2, & k \neq j. \end{cases} \quad (55)$$

Finally, we summarize the proposed low-complexity algorithm to solve the original optimization problem (37) in general multiuser case in Table II.

TABLE II
LOW-COMPLEXITY ALGORITHM FOR MULTI-USER FICIC

1. **Initialization:** Set $i = 0$, $R_{sum} = 0$, and $\bar{R}_{sum} = \sum_{k=1}^{K_P} \log(1 + \text{SINR}_k^{ub})$, where $\text{SINR}_k^{ub} = \frac{P_0 \|\mathbf{h}_{P_k}\|^2}{\sigma_n^2}$ is an upper bound of the SINR of PUE_k, which is achieved when the ICI disappears and only PUE_k is served by the PBS.
2. **Bisection Iteration:** At the i -th iteration, set $i \leftarrow i + 1$.
 - 1) Compute $R_{sum}^0 = \frac{R_{sum} + \bar{R}_{sum}}{2}$.
 - 2) Given R_{sum}^0 , compute $\{\lambda_k^*\}$ with the fixed-point iterative algorithm given by (47).
 - 3) Given $\{\lambda_k^*\}$, compute \mathbf{W}_f^* based on (50).
 - 4) Given R_{sum}^0 , $\{\lambda_k^*\}$ and \mathbf{W}_f^* , compute $\{\tilde{\mathbf{w}}_{d,k}^*\}$ and $\{p_k^*\}$ based on (52) and (54), respectively, then obtain $\mathbf{w}_{d,k}^* = \sqrt{p_k^*} \tilde{\mathbf{w}}_{d,k}^*$.
 - 5) Given \mathbf{W}_f^* and $\{\mathbf{w}_{d,k}^*\}$, compute the value of the objective function of problem (39), denoted by P_0^i .
 - 6) If $P_0^i \leq P_0$, let $\bar{R}_{sum} \leftarrow R_{sum}^0$, **otherwise**, let $\bar{R}_{sum} \leftarrow R_{sum}^0$.
3. **Repeat:** Iterate step 2 until the required accuracy is reached, i.e., $\bar{R}_{sum} - R_{sum} \leq \varepsilon$, where ε is a specific threshold.

V. GENERALIZATION TO WIDEBAND SYSTEMS

In previous sections the FICIC is optimized in narrowband systems, where all the channels are flat fading and therefore only single-tap forwarding precoder is designed. When considering wideband systems, frequency-selective fading channels should be taken into account and hence a multi-tap finite impulse response (FIR) forwarding precoder needs to be designed, which makes the design of the FICIC more involved.

In this section, we generalize the FICIC to OFDM systems. For the sake of notational simplicity, we consider single-user case, i.e., $K_M = 1$ and $K_P = 1$, which can be straightforwardly extended to multiuser orthogonal frequency division multi-accessing (OFDMA) systems due to the orthogonality between subcarriers. Following the previous narrowband-system definitions, let $\mathbf{h}_{Mk}(t)$ and $\mathbf{h}_{P_k}(t)$ denote the time-domain channels from the MBS and the PBS to PUE_k, $\mathbf{H}_{MP}(t)$ denote the time-domain channel from the MBS to the PBS, and $\mathbf{H}_{PP}(t)$ denote the time-domain self-interference channel of the FD PBS. The corresponding frequency-domain channels on the n -th subcarrier of the four channels are denoted by \mathbf{g}_{Mkn} , \mathbf{g}_{Pkn} , \mathbf{G}_{MPn} , and \mathbf{G}_{PPn} , respectively.

Consider that the PBS employs a FIR precoder $\mathbf{W}_f(t) = \sum_{l=0}^{L-1} \mathbf{W}_{fl} \delta(t - lT_s)$ to forward the listened ICI, where L is the order of the FIR precoder and T_s is the sampling interval. The selection of L needs to ensure that the delay spread of the equivalent channel for the forwarded ICI, $\mathbf{h}_{P_k}(t) \odot \mathbf{W}_f(t) \odot \mathbf{H}_{MP}(t) \odot \delta(t - \tau)$, does not exceed the cyclic prefix (CP) of the OFDM system in order to maintain orthogonality between subcarriers. Let $\bar{\mathbf{W}}_{fn}$ denote the frequency response of $\mathbf{W}_f(t)$ on the n -th subcarrier.

Then, following the same derivations in Section II for narrowband systems, we can obtain the transmit power of the PBS on the n -th subcarrier as

$$P_{out,n} = \frac{\text{tr}(\bar{\mathbf{W}}_{fn} \bar{\mathbf{g}}_{MPn} \bar{\mathbf{g}}_{MPn}^H \bar{\mathbf{W}}_{fn}^H) + \sigma_n^2 \text{tr}(\bar{\mathbf{W}}_{fn} \bar{\mathbf{W}}_{fn}^H) + \|\bar{\mathbf{w}}_{dn}\|^2}{1 - \sigma_e^2 \text{tr}(\bar{\mathbf{W}}_{fn} \bar{\mathbf{W}}_{fn}^H)}, \quad (56)$$

where $\bar{\mathbf{g}}_{MPn} = \mathbf{G}_{MPn}^H \bar{\mathbf{w}}_{Mn}$ is the equivalent frequency-domain channel from the MBS to the PBS, $\bar{\mathbf{w}}_{Mn}$ is the

precoder at the MBS for the MUE on the n -th subcarrier, $\bar{\mathbf{w}}_{dn}$ is the precoder at the PBS for the desired signal of PUE $_k$, and $\sigma_e^2 = \frac{\sigma_n^2}{P_{tr}} + 2\bar{\alpha}_{PP}(\mu_x + \mu_y)$ is the same as defined in (11). With (56), the total transmit power constraint for the PBS can be expressed as

$$P_{out} = \sum_{n=1}^N P_{out,n} \leq P_0. \quad (57)$$

Compared to the power constraint in (12) for narrowband system, which can be converted into a convex constraint for the precoders as shown in (14), the constraint in (57) for wideband system is much more complicated and is non-convex.

Similar to (17), we can obtain the SINR of PUE $_k$ on the n -th subcarrier as

$$\begin{aligned} \text{SINR}_{kn,FD} &= |\mathbf{g}_{Pkn}^H \bar{\mathbf{w}}_{dn}|^2 / (\|\bar{\mathbf{g}}_{Mkn} + \mathbf{g}_{Pkn}^H \bar{\mathbf{W}}_{fn} \bar{\mathbf{g}}_{MPn} e^{-jdn}\|^2 \\ &\quad + \|\mathbf{g}_{Pkn}^H \bar{\mathbf{W}}_{fn}\|^2 (P_{out,n} \sigma_e^2 + \sigma_n^2) + \sigma_n^2), \end{aligned} \quad (58)$$

where $\bar{\mathbf{g}}_{Mkn} \triangleq \bar{\mathbf{w}}_{Mn}^H \mathbf{g}_{Mkn}$ is the equivalent channel from the MBS to the PUE on the n -th subcarrier, and $P_{out,n}$ is given in (56), which is a function of the precoders $\bar{\mathbf{W}}_{fn}$ and $\bar{\mathbf{w}}_{dn}$.

Then, the wideband fICIC precoder optimization problem, aimed at maximizing the sum rate of PUE $_k$ over all subcarrier, can be formulated as

$$\max_{\mathbf{W}_f(t), \{\bar{\mathbf{w}}_{dn}\}} \sum_{n=1}^N \log(1 + \text{SINR}_{kn,FD}) \quad (59a)$$

$$s.t. [\bar{\mathbf{W}}_{f1}]_{ij}, \dots, [\bar{\mathbf{W}}_{fN}]_{ij}^T = \mathbf{F} [[\mathbf{W}_{f1}]_{ij}, \dots, [\mathbf{W}_{fL}]_{ij}]^T \quad (59b)$$

$$\sum_{n=1}^N P_{out,n} \leq P_0, P_{out,n} \geq 0, \forall n, \quad (59c)$$

where constraint (59b) restricts that the N frequency-domain precoders $\{\bar{\mathbf{W}}_{fn}\}$ are generated from the L -tap time-domain FIR precoder $\mathbf{W}_f(t)$, and $\mathbf{F} \in \mathbb{C}^{N \times L}$ is the matrix containing the first L columns of the $N \times N$ fast fourier transformation matrix.

Problem (59) is non-convex, whose global optimal solution is difficult to find. We can obtain a local optimal solution to the problem by using a gradient-based solution (specifically using the function `fmincon` of the optimization toolbox of MATLAB). Note that the direction of $\bar{\mathbf{w}}_{dn}$ only affects the power of desired signal in the numerator of $\text{SINR}_{kn,FD}$. Therefore, the direction of the optimal $\bar{\mathbf{w}}_{dn}$ can be obtained as $\frac{\bar{\mathbf{w}}_{dn}^*}{\|\bar{\mathbf{w}}_{dn}^*\|} = \frac{\mathbf{g}_{Pkn}}{\|\mathbf{g}_{Pkn}\|}$, with which the number of variables in problem (59) is reduced.

VI. PRACTICAL ISSUES

By now, we have introduced the concept of fICIC and optimized the associated precoders. In this section, we discuss some practical issues regarding the application of the fICIC.

A. Channel Acquisition

To apply the fICIC, the PBS needs to have the channels from the MBS to both the PBS and PUEs, i.e., \mathbf{H}_{MP} and \mathbf{h}_{Mk} , as well as the channel from the PBS to PUEs, \mathbf{h}_{Pk} , $\forall k$. First, the channel \mathbf{H}_{MP} can be directly estimated at the PBS by using the FD receive antennas to receive the downlink training signals sent from the MBS. Second, noting that in TDD systems channel reciprocity holds between the HD transceiver at the PBS and each PUE, the channel \mathbf{h}_{Pk} can be estimated at the PBS by using the HD receive antennas to receive the uplink training signals sent by PUE $_k$. Finally, the channel \mathbf{h}_{Mk} can be first estimated by PUE $_k$ and then fed back to the PBS, where digital or analog feedback schemes can be employed [18]. We will evaluate the performance of the fICIC under imperfect channel estimation and feedback in next section.

B. fICIC v.s. the HD Scheme

As we analyzed in Section III-B, with the fICIC, the FD PBS can adaptively switch between FD mode and HD mode by optimizing the precoders for transmitting desired signals and forwarding listened interference. For instance, when the ICI is very strong or the residual self-interference is very large, the fICIC will reduce to the HD scheme, as shown by (33) and (35). Therefore, with perfect channel information at the PBS, the proposed fICIC will always outperform the HD scheme given the same number of uplink and downlink RF chains and BB modules. Yet, to support the FD technique extra passive antennas (though cheap) and FD modules are required. When imperfect channels are considered at the PBS, the performance of the fICIC will decrease. Nevertheless, simulations in next section show that significant performance gain can be still achieved by the fICIC over the HD scheme with imperfect channels.

C. Self-interference Cancellation

Effective self-interference cancellation is crucial for the fICIC, where isolation of the transmit and FD receive antennas is an important approach [6]. Considering the transceiver structure of the FD PBS given in Fig. 1(a), where the FD receive antennas are only active in the downlink to receive the ICI from the MBS while not receiving the uplink signals from the PUEs, the FD receive antennas can be mounted far away from the transmit antennas, e.g., installing the FD receive antennas outside a building and transmit antennas inside, respectively. In this manner, the self-interference can be largely suppressed in general.

D. Joint Application with Existing ICIC Schemes

Existing ICIC schemes including eICIC and CoMP-CB operate at the MBS, while the fICIC can operate at each PBS individually. Therefore, the fICIC can be directly applied together with existing ICIC schemes. As will be shown in next section, a joint application of the fICIC and existing schemes can achieve evident performance improvement.

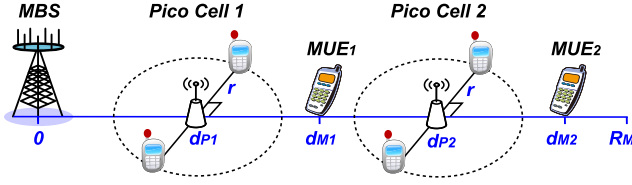


Fig. 2. Network layout for simulations, where two PBSs are deployed in the macro cell covered by the MBS.

VII. SIMULATION RESULTS

In this section we verify our analytical results and evaluate the performance of the proposed fICIC scheme via simulations.

A. Simulation Setups

For all simulations, unless otherwise specified, the parameters in this subsection are used. The considered HetNet layout is shown in Fig. 2, where the MBS is located at the center of the macro cell, PBS₁ and PBS₂ are respectively located at $(d_{P1}, 0)$ and $(d_{P2}, 0)$, the MBS serves two MUEs located at $(d_{M1}, 0)$ and $(d_{M2}, 0)$, and PBS_k serves two PUEs located at (d_{Pk}, r) and $(d_{Pk}, -r)$, respectively. We set the radius of the macro cell R_M as 500 m, $d_{P1} = 60$ m, $d_{P2} = 180$ m, $r = 40$ m, $d_{M1} = 120$ m, and $d_{M2} = 240$ m. Since we have shown in the analytical results that the performance of the fICIC depends on the strength of the ICI, the positions of the two PBSs, d_{P1} and d_{P2} , are selected to reflect the cases of strong and weak ICI, respectively. We will also evaluate the performance of the fICIC with random PBS placements later.

The MBS transmits with $M = 4$ antennas and the power of $P_M = 46$ dBm, each PBS has $N_r = 2$ FD receive antennas and transmits with $N_t = 2$ antennas and the maximal power of $P_0 = 30$ dBm, and each UE (including MUE and PUE) has one receive antenna. The path loss is set as $128.1 + 37.6 \log_{10} d$ for macro cell and $140.7 + 36.7 \log_{10} d$ for pico cell, respectively, where d is the distance in km [19]. A penetration loss of 20 dB is considered for the channels to PUEs. We model the interference from the MBSs in adjacent macro cells and the surrounding PBSs as noises. Define the average receive SNR of a MUE located at the cell edge as SNR_{edge} , then the noise variance σ_n^2 can be obtained as $\sigma_n^2 = P_M - (128.1 + 37.6 \log_{10} R_M) - \text{SNR}_{\text{edge}}$ in dBm. To evaluate the impact of imperfect self-interference cancellation for FD, we define the signal to self-interference ratio as $\text{SIR}_{\text{self}} = P_0 - \sigma_I^2$ in dB in order to reflect the level of self-interference cancellation. The Rayleigh flat small-scale fading channels are considered. The fairness factors are set as $\alpha_k = \frac{1}{K_P}$ in multi-user case. All the results are averaged over 1000 channel realizations.

B. Narrowband Single-user Case

We first simulate the single-user case, where the MBS serves one MUE, each PBS serves one PUE, and each PBS has $N_r = 1$ FD receive antenna.

Under the assumption of perfect self-interference cancellation, the average sum rate achieved by the fICIC is depicted

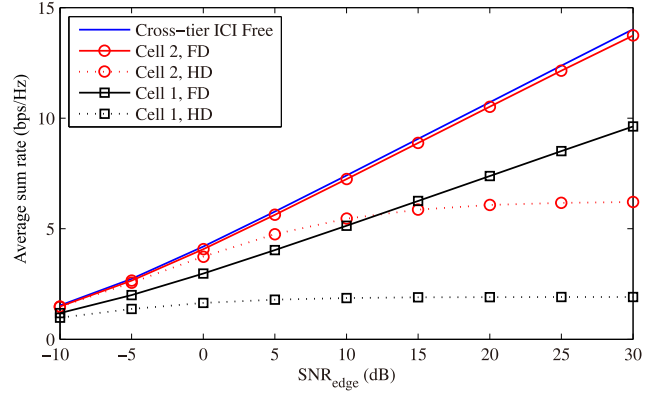


Fig. 3. Average sum rate versus SNR_{edge} , where perfect self-interference cancellation and perfect channels are considered, $K_M = 1$, $K_P = 1$, and $N_r = 1$.

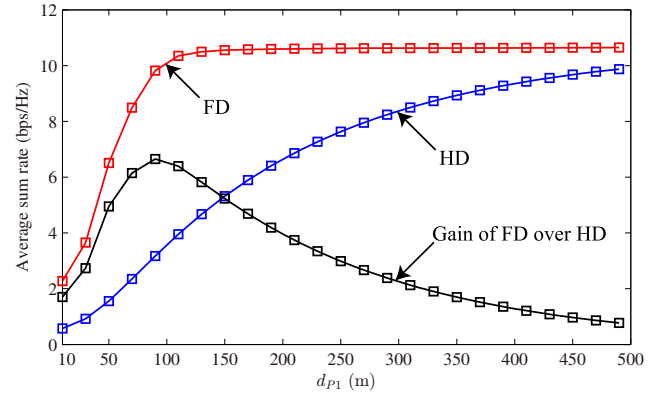


Fig. 4. Average sum rate versus d_{P1} for different PBS placements, where perfect self-interference cancellation and perfect channels are considered, $K_M = 1$, $K_P = 1$, and $N_r = 1$.

in Fig. 3. For comparison, the performance of the HD scheme, given by $\mathcal{E}\{\log(1 + \text{SINR}_{k,HD}^*)\}$, and the performance in ICI free case, given by $\mathcal{E}\{\log(1 + \frac{P_0 \|\mathbf{h}_{Pk}\|^2}{\sigma_n^2})\}$, are also presented. First, we can see that all the schemes perform closely in low SNR regime, where the system operates in noise-limited scenario. With the increase of SNR, the performance floor appears for the HD scheme, which is caused by the ICI from the MBS. The performance of Cell 2 is better than Cell 1 due to $d_{P2} > d_{P1}$ that leads to weaker ICI. The proposed fICIC exhibits a noticeable performance gain over the HD scheme. For weak ICI case, e.g., Cell 2, as we analyzed, the fICIC can thoroughly eliminate the ICI in a high probability (considering the randomness of small-scale channels), and thus the performance gap between the fICIC and the ICI free case is very small.

In Fig. 4, we evaluate the impact of the strength of the ICI on the performance of the fICIC by simulating a single PBS with different positions d_{P1} . The performance achieved by the fICIC and the HD scheme as well as the performance gain of the fICIC are depicted. It is shown that the performance of the HD scheme increases slowly with the reduction of ICI, i.e., the increase of d_{P1} , while the performance of the fICIC first increases fast and then keeps nearly constant. It implies that the fICIC can make a large area of the macro cell, e.g.,

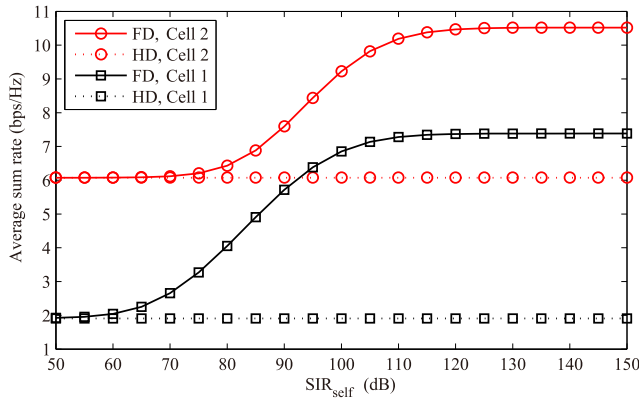


Fig. 5. Average sum rate versus SIR_{self} with imperfect self-interference cancellation, where perfect channels are considered, $\text{SNR}_{\text{edge}} = 20$ dB, $K_M = 1$, $K_P = 1$, and $N_r = 1$.

from 150 m to 500 m, experience very weak ICI. Moreover, we can see that the fICIC performs close to the HD scheme at large d_{P1} because the ICI is very weak and has negligible impact on the performance in this case. Further recalling that the fICIC will degenerate to the HD scheme when the ICI is very strong as analyzed in Section III-B, we can understand that the gain of the fICIC over the HD scheme first increases and then decreases with d_{P1} . However, we can still observe an evident performance improvement of nearly 300% even when $d_{P1} = 10$ m, in which case the average power of the ICI is 10.4 dB stronger than that of the desired signals.

Figure 5 shows the performance of the fICIC as a function of SIR_{self} , where imperfect self-interference cancellation is considered. First, it can be seen that the fICIC reduces to the HD scheme for low SIR_{self} , which agrees with our previous analysis. Second, the fICIC outperforms the HD scheme when $\text{SIR}_{\text{self}} \geq 60$ dB for Cell 1 but 75 dB for Cell 2. This is because PBS_1 is closer to the MBS and hence can listen a stronger ICI, which can relax the requirement of self-interference cancellation for FD. Moreover, it is shown that the self-interference cancellation of 110 dB is sufficient for the fICIC, which is practically possible because on one hand existing work has reported 90 dB self-interference cancellation even for closely placed transmit and receive antennas [6], and on the other hand, as we discussed before, in our case the transmit antennas and FD receive antennas can be separated far apart, leading to further reduction of the self-interference, e.g., with an additional 20 dB penetration loss.

C. Narrowband Multi-user Case

In this subsection the multi-user case is simulated, where each PBS serves two PUEs. We compare the performance of the fICIC with the HD scheme, a successive interference cancellation (SIC) based non-linear HD scheme (denoted by HD-SIC) [20], the ABS-based eICIC and CoMP-CB, and also evaluate the performance of the combinations of the fICIC with eICIC and CoMP-CB. To study the effectiveness of the seven schemes for cancelling different-strength ICI, we fix the location of Cell 2 and adjust the location of Cell 1 to generate different interference to noise ratio (INR) for PUEs in Cell

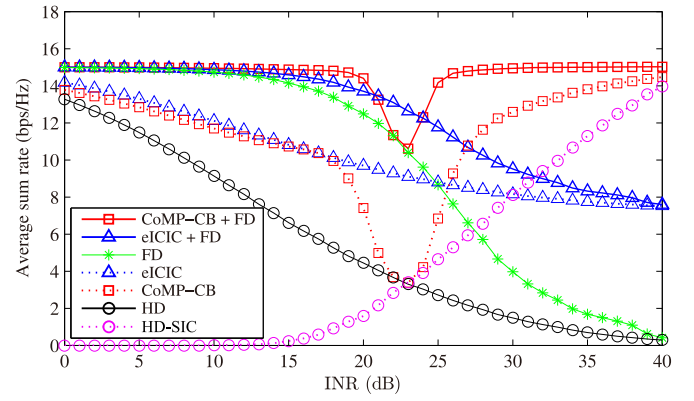


Fig. 6. Average sum rate of PUEs in Cell 1 achieved by relevant schemes versus INR, where perfect self-interference cancellation and perfect channels are considered, $\text{SNR}_{\text{edge}} = 20$ dB, $N_r = 2$, $K_M = 1$, and $K_P = 2$.

1, where $\text{INR} \triangleq \frac{\mathcal{E}\{\|\mathbf{h}_{Mk}\|^2\}}{\sigma_p^2}$. To enable the HD-SIC scheme, we consider that the MBS serves a single MUE because each PUE has only one antenna so that only one interference signal can be decoded and cancelled.

In the simulations, the proposed low-complexity algorithm given in Table II is employed to obtain the performance of both the fICIC and the HD scheme, where \mathbf{W}_f is set as zero for the HD scheme. For the HD-SIC, since the ICI needs to be decoded first by regarding the desired signal from the PBS as interference, the PBS may not transmit with its maximal power. Given the data rate of the MUE as 4 bps/Hz, we employ exhaustive searching to find the maximal feasible transmit power of the PBS under the ICI decoding constraint, where for any given transmit power the precoder of the PBS is computed the same as the HD scheme. For the ABS-based eICIC, it is considered that the MBS mutes in half time to provide ICI-free environment to the PUEs. For CoMP-CB, the signal-to-leakage-plus-noise ratio (SLNR) based precoder [21] is employed. The results are depicted in Fig. 6.

Compared with the fICIC, we can see that the eICIC achieves worse performance when the ICI is not strong, say $\text{INR} < 25$ dB, because the fICIC can effectively cancel the weak-medium level of ICI and allow the PUEs to use all the time-frequency resources, while with the eICIC the PUEs only uses half resources. When the ICI is strong, as we analyzed before, the fICIC is not effective, and the eICIC shows large performance gain. By combining eICIC with fICIC, the two schemes supplement each other and achieve much better performance.

For CoMP-CB, we can see that it is superior to the fICIC when the ICI is strong, say $\text{INR} \geq 26$ dB, but inferior for weak-medium ICI. This can be explained as follows. In the simulated case, the MBS has only four antennas, which are not adequate to serve one MUE and suppress the ICI for four PUEs simultaneously. As a result, when the INR of Cell 1 is smaller than Cell 2, the SLNR-based CoMP-CB only suppresses the stronger ICI to the PUEs in Cell 2, while the ICI to PUEs in Cell 1 will be mitigated only for large INR (this also explains why the performance of Cell 1 with pure CoMP-CB or the combination of CoMP-CB and fICIC

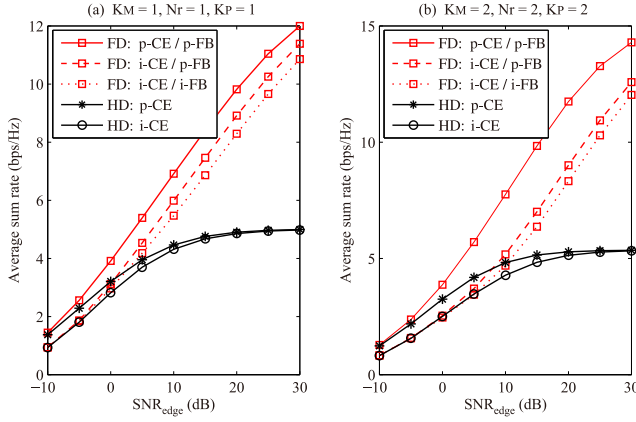


Fig. 7. Average sum rate of PUEs in Cell 1 versus SNR_{edge} , where PBS_1 is uniformly placed within the area suffering from strong ICI with $d_{P1} \sim U([50, 250])$ m, $\text{SIR}_{\text{self}} = 110$ dB, $N_r = 1, 2$, $K_M = 1, 2$, and $K_P = 1, 2$. In the legends, “p-CE” and “i-CE” denote perfect and imperfect channel estimation, respectively, and “p-FB” and “i-FB” denote perfect and imperfect feedback, respectively.

first decreases and then increases with the growth of INR.). By contrast, the fICIC is effective to suppress weak-medium ICI but not for strong ICI. Therefore, CoMP-CB and fICIC perform better in different ICI cases. Since CoMP-CB has no enough antenna resources to cancel weak-medium ICI in the simulation, the combination of CoMP-CB and fICIC has negligible gain over the pure fICIC for weak-medium INR. However, evident performance improvement is achieved by their combination for strong ICI, because the strong ICI can be first suppressed into weak-medium ICI by CoMP-CB and then further effectively cancelled by the fICIC.

For the HD-SIC, we can observe that it is not always feasible when the ICI is not strong, e.g., $\text{INR} < 15$ dB. With the increase of ICI, the interference signal becomes decodable and more power can be transmitted by the PBS, which results in the performance improvement as expected. It can be seen that the HD-SIC outperforms the fICIC for strong ICI, which inspires us to investigate the SIC-based fICIC in future work, where the FD PBS forwards the listened ICI to further enhance the ICI at the PUE so as to extend the feasible region of SIC.

Finally, in Fig. 7 the performance of the fICIC under imperfect self-interference cancellation, imperfect channel estimation and practical channel feedback is evaluated, where only Cell 1 is considered and PBS_1 is randomly placed within the area experiencing strong ICI, specifically with d_{P1} following uniform distribution between $[50, 250]$ m. As discussed before, the channel \mathbf{H}_{MP} can be directly estimated at the PBS, the channel \mathbf{h}_{Pk} can be obtained at the PBS by estimating the uplink channel based on channel reciprocity, and the channel \mathbf{h}_{Mk} can be first estimated at the PUE_k and then fed back to the PBS. In simulations, we employ linear minimum mean-squared error estimator to estimate \mathbf{H}_{MP} , \mathbf{h}_{Pk} and \mathbf{h}_{Mk} , and use analog feedback [18] to send back the estimate of \mathbf{h}_{Mk} to the PBS, where the transmit power of PUE_k is set as 23 dBm. Figs. 7(a) and 7(b) show the results in single-user and multi-user cases, respectively. In both cases we can see the performance degradation of the fICIC caused by imperfect channel estimation and imperfect

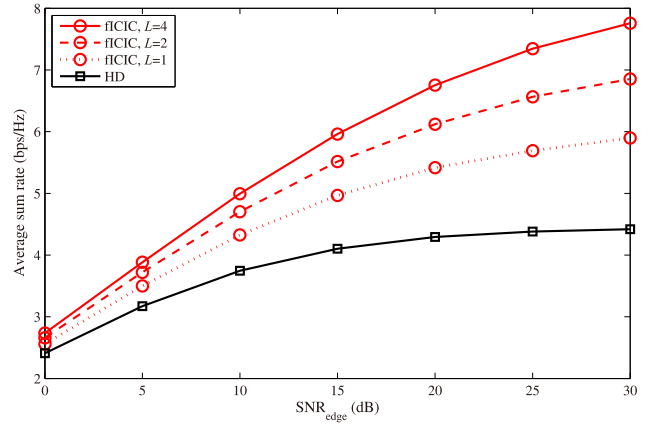


Fig. 8. Average sum rate of the PUE over $N = 25$ subcarriers in Cell 1 versus SNR_{edge} , where $\text{SIR}_{\text{self}} = 110$ dB, $K_M = 1$, and $K_P = 1$.

feedback. However, compared with the HD scheme, significant performance gain achieved by the fICIC can be still observed, even when imperfect self-interference cancellation, imperfect channel estimation and practical analog feedback are taken into account.

D. Wideband Single-user Case

In this subsection we evaluate the performance of the fICIC in wideband systems. We consider a LTE system with 5 MHz bandwidth, where the sampling interval is $T_s = 0.13 \mu\text{s}$ [22]. The small-scale channels are generated based on WINNER II clustered delay line model [23]. Specifically, the channels from the MBS to the PBS, \mathbf{H}_{MP} , use the typical urban macro-cell line of sight (LoS) model considering that the receive antennas of the PBS can be mounted outside a building as discussed before, the channels from the MBS to PUEs, \mathbf{h}_{Mk} , use the typical urban macro-cell non-LoS model, and the channels from the PBS to PUEs, \mathbf{h}_{Pk} , use the typical urban micro-cell NLoS model. After sampling the multipath channels with the considered bandwidth, we obtain the maximal delay spread of \mathbf{H}_{MP} and \mathbf{h}_{Pk} as three and six samples, respectively. We consider the processing delay of the FD PBS as 4 samples, i.e., $\tau = 0.52 \mu\text{s}$. Since the CP of the LTE system is $4.7 \mu\text{s}$ [22], i.e., 36 samples, we can obtain the maximal order of the FIR forwarding precoder $\mathbf{W}_f(t)$ as 23, i.e., $L \leq 23$. Note that the maximum value of L can be even larger if the extended CP with $16.7 \mu\text{s}$ is considered.

Considering that the complexity of solving problem (59) increases rapidly with the number of subcarriers and also considering the channel correlation in frequency domain, in the simulations we treat each resource block (RB) as a subcarrier and then $N = 25$ subcarriers are simulated corresponding to the total 25 RBs of the 5 MHz LTE system [22]. In addition, we select the order of \mathbf{W}_f as $L = 1, 2, 4$ to speed up the simulations. In Fig. 8, the average sum rate of Cell 1 with $d_{P1} = 60$ m achieved by the fICIC is depicted, where imperfect self-interference cancellation, imperfect channel estimation and practical channel feedback as considered in Fig. 7 are taken into account. The compared HD scheme is obtained from problem (59) by setting \mathbf{W}_f as zeros.

We can see from Fig. 8 that the performance of the wide-band fICIC improves with the increase of L , and an evident performance gain over the HD scheme can be observed when $L = 4$. More significant performance gain can be expected when larger L is used for the fICIC, to achieve which more efficient low-complexity algorithms to problem (59) need to be studied in future work.

VIII. CONCLUSIONS

In this paper we proposed to eliminate the cross-tier ICI in HetNets by using FD technique at the PBS. We derived a FD assisted ICI cancellation (fICIC) scheme, with which the ICI can be mitigated by merely designing the precoders at PBSs without relying on the participation of the MBS. We first investigated the narrowband single-user case to gain some insight into the behavior of the fICIC, where we found closed-form solution of the optimal fICIC, and analyzed its asymptotical performance in ICI-dominated scenario. We then studied the general narrowband multi-user case, and devised a low-complexity algorithm to find the optimal fICIC scheme that maximizes the downlink sum rate of PUEs subject to the fairness constraint. Finally, we generalized the fICIC to wideband systems and discussed the practical issues regarding the application of the fICIC. Simulations validated the analytical results. Compared with the traditional HD scheme, the fICIC exhibits significant performance gain even when imperfect self-interference cancellation and imperfect channel information are taken into account. By combining the fICIC with eICIC or CoMP-CB, the ICI with various levels can be effectively eliminated.

APPENDIX A DERIVATION OF TRANSMIT POWER P_{out}

By substituting (8) into (9), we can obtain

$$\mathbf{x}_p[t] = \mathbf{W}_f(\bar{\mathbf{H}}_{MP}^H \mathbf{s}_M[t-\tau] - \mathbf{E}_{PP}^H[t-\tau] \mathbf{x}_p[t-\tau] + \mathbf{H}_{PP}^H \mathbf{z}_x[t-\tau] + \mathbf{n}_p[t-\tau] + \mathbf{z}_y[t-\tau]) + \sum_{k=1}^{K_P} \mathbf{w}_{d,k} s_{p,k}[t]. \quad (\text{A.1})$$

Since the terms at the right-hand side of (A.1) are independent, we can obtain P_{out} with (10) by taking the expectations over each term, yielding

$$P_{out} = \text{tr}(\mathbf{W}_f \bar{\mathbf{H}}_{MP}^H \bar{\mathbf{H}}_{MP} \mathbf{W}_f^H) + \sigma_n^2 \text{tr}(\mathbf{W}_f \mathbf{W}_f^H) + \sum_{k=1}^{K_P} \|\mathbf{w}_{d,k}\|^2 + \text{tr}(\mathbf{W}_f \Sigma_1 \mathbf{W}_f^H) + \text{tr}(\mathbf{W}_f \Sigma_2 \mathbf{W}_f^H) + \text{tr}(\mathbf{W}_f \Sigma_3 \mathbf{W}_f^H), \quad (\text{A.2})$$

where $\Sigma_1 \triangleq \mathcal{E}_{\mathbf{x}_p, \mathbf{E}_{PP}, \mathbf{H}_{PP}}\{\mathbf{E}_{PP}^H[t-\tau] \mathbf{x}_p[t-\tau] \mathbf{x}_p^H[t-\tau] \mathbf{E}_{PP}[t-\tau]\}$, $\Sigma_2 \triangleq \mathcal{E}_{\mathbf{H}_{PP}}\{\mathbf{H}_{PP}^H \mu_x \text{diag}(\Phi_x) \mathbf{H}_{PP}\}$, and $\Sigma_3 \triangleq \mathcal{E}_{\mathbf{H}_{PP}}\{\mu_y \text{diag}(\Phi_y)\}$, which can be derived as follows.

Denoting $\mathbf{x}_p = [x_{p1}, \dots, x_{pN_t}]^T$, we can rewrite Σ_1 as

$$\begin{aligned} \Sigma_1 &= \mathcal{E}_{\mathbf{x}_p, \mathbf{E}_{PP}, \mathbf{H}_{PP}}\left\{\sum_{i=1}^{N_t} |x_{pi}[t-\tau]|^2 \mathbf{e}_{PP,i} \mathbf{e}_{PP,i}^H\right\} \\ &\stackrel{(a)}{=} \mathcal{E}_{\mathbf{H}_{PP}}\left\{\sum_{i=1}^{N_t} [\Phi_x]_{ii} \tilde{\Phi}_{e,i}\right\} \\ &\stackrel{(b)}{=} \mathcal{E}_{\mathbf{H}_{PP}}\left\{\sum_{i=1}^{N_t} [\Phi_x]_{ii} \left(\mathbf{H}_{PP}^H \sum_{j=1}^{N_t} |c_{ij}|^2 \mu_x \text{diag}(\mathbf{c}_j \mathbf{c}_j^H) \mathbf{H}_{PP} + \frac{1}{P_{tr}} (1 + \mu_y) \sigma_n^2 \mathbf{I}_{N_r} + \sum_{j=1}^{N_t} |c_{ij}|^2 \mu_y \cdot \right. \right. \\ &\quad \left. \left. \text{diag}(\mathbf{H}_{PP}^H (\mathbf{c}_j \mathbf{c}_j^H + \mu_x \text{diag}(\mathbf{c}_j \mathbf{c}_j^H)) \mathbf{H}_{PP})\right)\right\} \\ &\stackrel{(c)}{=} \sum_{i=1}^{N_t} [\Phi_x]_{ii} \left(\sum_{j=1}^{N_t} |c_{ij}|^2 |\mu_x \bar{\alpha}_{PP} \text{tr}(\text{diag}(\mathbf{c}_j \mathbf{c}_j^H)) \mathbf{I}_{N_r} + \frac{1}{P_{tr}} (1 + \mu_y) \sigma_n^2 \mathbf{I}_{N_r} + \sum_{j=1}^{N_t} |c_{ij}|^2 \mu_y \bar{\alpha}_{PP} (\text{tr}(\mathbf{c}_j \mathbf{c}_j^H) \mathbf{I}_{N_r} + \mu_x \text{tr}(\text{diag}(\mathbf{c}_j \mathbf{c}_j^H) \mathbf{I}_{N_r}))\right) \\ &\stackrel{(d)}{=} \left(\frac{(1 + \mu_y) \sigma_n^2}{P_{tr}} + \bar{\alpha}_{PP} (\mu_x + \mu_y + \mu_x \mu_y)\right) \text{tr}(\Phi_x) \mathbf{I}_{N_r}, \quad (\text{A.3}) \end{aligned}$$

where the expectations over \mathbf{x}_p and \mathbf{E}_{PP} are taken in step (a), $[\Phi_x]_{ii}$ denotes the i -th diagonal element of Φ_x , step (b) comes from (7), step (c) takes the expectation over \mathbf{H}_{PP} considering $\text{vec}(\mathbf{H}_{PP}) \sim \mathcal{CN}(\mathbf{0}_{N_r N_t}, \bar{\alpha}_{PP} \mathbf{I}_{N_r N_t})$, and step (d) comes from the fact that $\mathbf{C} \mathbf{C}^H = \mathbf{I}_{N_t}$ such that $\text{tr}(\mathbf{c}_i \mathbf{c}_i^H) = \text{tr}(\text{diag}(\mathbf{c}_i \mathbf{c}_i^H)) = \sum_{j=1}^{N_t} |c_{ij}|^2 = 1$.

The term Σ_2 can be obtained as

$$\Sigma_2 = \mathcal{E}_{\mathbf{H}_{PP}}\{\mathbf{H}_{PP}^H \mu_x \text{diag}(\Phi_x) \mathbf{H}_{PP}\} = \bar{\alpha}_{PP} \mu_x \text{tr}(\Phi_x) \mathbf{I}_{N_r}. \quad (\text{A.4})$$

With (3), the term Σ_3 can be obtained as

$$\begin{aligned} \Sigma_3 &= \mathcal{E}_{\mathbf{H}_{PP}}\{\mu_y \text{diag}(\Phi_y)\} \\ &= \mathcal{E}_{\mathbf{H}_{PP}}\left\{\mu_y \text{diag}(\bar{\mathbf{H}}_{MP}^H \bar{\mathbf{H}}_{MP} + \mathbf{H}_{PP}^H (\Phi_x + \mu_x \text{diag}(\Phi_x)) \mathbf{H}_{PP} + \sigma_n^2 \mathbf{I}_{N_r})\right\} \\ &= \mu_y \left(\text{diag}(\bar{\mathbf{H}}_{MP}^H \bar{\mathbf{H}}_{MP}) + \bar{\alpha}_{PP} (1 + \mu_x) \text{tr}(\Phi_x) \mathbf{I}_{N_r} + \sigma_n^2 \mathbf{I}_{N_r}\right). \end{aligned} \quad (\text{A.5})$$

Substituting (A.3), (A.4) and (A.5) into (A.2) and noting that $\text{tr}(\Phi_x) = P_{out}$, we can obtain

$$\begin{aligned} P_{out} &= \text{tr}(\mathbf{W}_f (\bar{\mathbf{H}}_{MP}^H \bar{\mathbf{H}}_{MP} + \mu_y \text{diag}(\bar{\mathbf{H}}_{MP}^H \bar{\mathbf{H}}_{MP})) \mathbf{W}_f^H) + \sum_{k=1}^{K_P} \|\mathbf{w}_{d,k}\|^2 + \left((1 + \mu_y) \sigma_n^2 + \left(\frac{(1 + \mu_y) \sigma_n^2}{P_{tr}} + 2\bar{\alpha}_{PP} (\mu_x + \mu_y + \mu_x \mu_y)\right) P_{out}\right) \text{tr}(\mathbf{W}_f \mathbf{W}_f^H) \\ &\approx \text{tr}(\mathbf{W}_f \bar{\mathbf{H}}_{MP}^H \bar{\mathbf{H}}_{MP} \mathbf{W}_f^H) + \sum_{k=1}^{K_P} \|\mathbf{w}_{d,k}\|^2 + \left(\sigma_n^2 + \left(\frac{\sigma_n^2}{P_{tr}} + 2\bar{\alpha}_{PP} (\mu_x + \mu_y)\right) P_{out}\right) \text{tr}(\mathbf{W}_f \mathbf{W}_f^H), \quad (\text{A.6}) \end{aligned}$$

where the approximation follows from $\mu_x \ll 1$ and $\mu_y \ll 1$ as in [8].

APPENDIX B PROOF OF LEMMA 1

To prove this lemma, we rewrite problem (21) by expressing \mathbf{W}_f with its vectorization, denoted by $\bar{\mathbf{w}}_f = \text{vec}(\mathbf{W}_f)$, as

follows.

$$\begin{aligned} \max_{\bar{\mathbf{w}}_f, \mathbf{w}_{d,k}} \quad & \frac{|\mathbf{h}_{P_k}^H \mathbf{w}_{d,k}|^2}{\Omega'_k} \quad (\text{B.1a}) \\ \text{s.t.} \quad & \|(\bar{\mathbf{h}}_{MP}^T \otimes \mathbf{I}_{N_t}) \bar{\mathbf{w}}_f\|^2 + (P_0 \sigma_e^2 + \sigma_n^2) \|\bar{\mathbf{w}}_f\|^2 + \|\mathbf{w}_{d,k}\|^2 \leq P_0, \quad (\text{B.1b}) \end{aligned}$$

where $\Omega'_k \triangleq |\bar{h}_{Mk}^* + e^{-j\phi} (\bar{\mathbf{h}}_{MP}^T \otimes \mathbf{h}_{P_k}^H) \bar{\mathbf{w}}_f|^2 + \|(\mathbf{I}_{N_r} \otimes \mathbf{h}_{P_k}^H) \bar{\mathbf{w}}_f\|^2 (P_0 \sigma_e^2 + \sigma_n^2) + \sigma_n^2$, and the property $\text{vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A}) \text{vec}(\mathbf{X})$ is used.

Based on the KKT condition, we can obtain the optimal solution of $\bar{\mathbf{w}}_f$ as

$$\begin{aligned} \bar{\mathbf{w}}_f^* = & -\bar{h}_{Mk}^* e^{j\phi} \left(\frac{\text{SINR}_{k,FD}}{\Delta_1} (\bar{\mathbf{h}}_{MP}^T \otimes \mathbf{h}_{P_k}^H)^H (\bar{\mathbf{h}}_{MP}^T \otimes \mathbf{h}_{P_k}^H) \right. \\ & + \frac{\text{SINR}_{k,FD} \Delta_2}{\Delta_1} (\mathbf{I}_{N_r} \otimes \mathbf{h}_{P_k}^H)^H (\mathbf{I}_{N_r} \otimes \mathbf{h}_{P_k}^H) + \lambda (\bar{\mathbf{h}}_{MP}^T \otimes \mathbf{I}_{N_t})^H \\ & \left. (\bar{\mathbf{h}}_{MP}^T \otimes \mathbf{I}_{N_t}) + \lambda \Delta_2 \mathbf{I}_{N_t N_r} \right)^{-1} (\bar{\mathbf{h}}_{MP}^T \otimes \mathbf{h}_{P_k}^H)^H, \quad (\text{B.2}) \end{aligned}$$

where $\Delta_1 = |\bar{h}_{Mk}^* + e^{-j\phi} (\bar{\mathbf{h}}_{MP}^T \otimes \mathbf{h}_{P_k}^H) \bar{\mathbf{w}}_f|^2 + \|(\mathbf{I}_{N_r} \otimes \mathbf{h}_{P_k}^H) \bar{\mathbf{w}}_f\|^2 (P_0 \sigma_e^2 + \sigma_n^2) + \sigma_n^2$, $\Delta_2 = P_0 \sigma_e^2 + \sigma_n^2$, and λ is the lagrangian multiplier.

By applying the properties of Kronecker product and after some regular manipulations, we can simplify (B.2) as

$$\begin{aligned} \bar{\mathbf{w}}_f^* = & -\bar{h}_{Mk}^* e^{j\phi} \left(((\bar{\mathbf{h}}_{MP} \bar{\mathbf{h}}_{MP}^H)^T + \Delta_2 \mathbf{I}_{N_r})^{-1} (\bar{\mathbf{h}}_{MP}^H)^T \right) \\ & \otimes \left(\left(\frac{\text{SINR}_{k,FD}}{\Delta_1} \mathbf{h}_{P_k} \mathbf{h}_{P_k}^H + \lambda \mathbf{I}_{N_t} \right)^{-1} \mathbf{h}_{P_k} \right). \quad (\text{B.3}) \end{aligned}$$

Then based on the matrix inversion lemma, we can further rewrite (B.3) as

$$\begin{aligned} \bar{\mathbf{w}}_f^* = & \frac{-\bar{h}_{Mk}^* e^{j\phi} \Delta_1 (\bar{\mathbf{h}}_{MP}^H)^T \otimes \mathbf{h}_{P_k}}{(\text{SINR}_{k,FD} \|\mathbf{h}_{P_k}\|^2 + \lambda \Delta_1) (\|\mathbf{h}_{MP}\|^2 + \Delta_2)} \\ \triangleq & -\bar{h}_{Mk}^* e^{j\phi} \cdot \beta \cdot (\bar{\mathbf{h}}_{MP}^H)^T \otimes \mathbf{h}_{P_k}, \quad (\text{B.4}) \end{aligned}$$

where $\beta = \frac{\Delta_1}{(\text{SINR}_{k,FD} \|\mathbf{h}_{P_k}\|^2 + \lambda \Delta_1) (\|\mathbf{h}_{MP}\|^2 + \Delta_2)}$ is a positive scalar.

According to (B.4), we can recover the optimal \mathbf{W}_f^* from $\bar{\mathbf{w}}_f^*$ as (24), which is of rank 1.

APPENDIX C

PROOF OF PROPOSITION 1

By defining $\nu \triangleq |\bar{h}_{Mk}| \|\mathbf{h}_{P_k}\| \|\bar{\mathbf{h}}_{MP}\| \beta = \frac{C}{2} \beta$, problem (25) can be rewritten as

$$\max_{\nu} f(\nu) \triangleq \frac{A - B\nu^2}{B\nu^2 - C\nu + D} \quad (\text{C.1a})$$

$$\text{s.t. } \nu^2 \leq \frac{A}{B} \quad (\text{C.1b})$$

$$\nu \geq 0, \quad (\text{C.1c})$$

where the parameters A, B, C and D are defined in Proposition 1.

Since the numerator of the objective function (C.1a) is concave, the denominator is convex, and both are differential, we know that the objective function of problem (C.1) is quasi-concave [14]. This suggests that the global optimal solution

to problem (C.1), denoted by $\bar{\nu}$, needs to satisfy the following KKT conditions,

$$-\nabla f(\bar{\nu}) + 2u\bar{\nu} - v = 0 \quad (\text{C.2a})$$

$$u(\bar{\nu}^2 - \frac{A}{B}) = 0, \quad u \geq 0 \quad (\text{C.2b})$$

$$v\bar{\nu} = 0, \quad v \geq 0, \quad (\text{C.2c})$$

where u and v are the lagrangian multipliers.

We next show that $u = 0$ and $v = 0$ for problem (C.1). First, we find that $\bar{\nu}^2 < \frac{A}{B}$ must hold for the constraint (C.1b), otherwise, when $\bar{\nu}^2 = \frac{A}{B}$, the objective function will be zero that is obviously non-optimal. Therefore, from (C.2b) we have $u = 0$. Second, the objective function satisfies $\nabla f(0) > 0$, which means that we can always find a small $\epsilon > 0$ making $f(\epsilon) > f(0)$, i.e., $\bar{\nu} > 0$ must hold. Therefore, based on (C.2c) we have $v = 0$.

Then the condition (C.2a) for $\bar{\nu}$ can be simplified as

$$2B\bar{\nu} - \frac{(B\bar{\nu}^2 - A)(2B\bar{\nu} - C)}{B\bar{\nu}^2 - C\bar{\nu} + D} = 0. \quad (\text{C.3})$$

From (C.3), we can derive the optimal value of the objective function as

$$f(\bar{\nu}) = \frac{2B\bar{\nu}}{C - 2B\bar{\nu}} = \frac{1}{\frac{C}{2B\bar{\nu}} - 1}. \quad (\text{C.4})$$

The condition (C.3) can be further expressed as a quadratic equation with respect to $\bar{\nu}$, which is

$$g(\bar{\nu}) \triangleq BC\bar{\nu}^2 - 2B(A + D)\bar{\nu} + AC = 0, \quad (\text{C.5})$$

whose two solutions can be obtained as

$$\bar{\nu} = \frac{A + D \pm \sqrt{(A + D)^2 - \frac{AC^2}{B}}}{C}. \quad (\text{C.6})$$

The feasibility of the two solutions is examined as follows.

Recalling that $0 \leq \bar{\nu} \leq \sqrt{\frac{A}{B}}$ from (C.1b) and (C.1c), we can see that $g(0) = AC > 0$ and $g(\sqrt{\frac{A}{B}}) < 0$ because

$$\begin{aligned} g(\sqrt{\frac{A}{B}}) &= 2AC - 2(A + D)\sqrt{AB} \\ &< 4P_0 |\bar{h}_{Mk}| \|\bar{\mathbf{h}}_{MP}\| \|\mathbf{h}_{P_k}\|^3 - 2(P_0 \|\mathbf{h}_{P_k}\|^2 + |\bar{h}_{Mk}|^2) \cdot \\ &\quad \sqrt{P_0 \|\bar{\mathbf{h}}_{MP}\|^2 \|\mathbf{h}_{P_k}\|^4} \\ &= -2\sqrt{P_0} \|\bar{\mathbf{h}}_{MP}\| \|\mathbf{h}_{P_k}\|^2 (|\bar{h}_{Mk}| - \sqrt{P_0} \|\mathbf{h}_{P_k}\|)^2, \quad (\text{C.7}) \end{aligned}$$

where the inequality is obtained by setting $\sigma_e^2 = 0$ and $\sigma_n^2 = 0$ in the definitions of B and D .

The results indicate that equation (C.5) has one and only one solution within $[0, \sqrt{\frac{A}{B}}]$, which is the smaller one in (C.6). Finally, recalling that $\nu = \frac{C}{2} \beta$, we can obtain the optimal β as shown after (26).

APPENDIX D

PROOF OF PROPOSITION 2

By defining $\mathbf{A} = \mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j \mathbf{h}_{P_j} \mathbf{h}_{P_j}^H$ and $\mathbf{B} = \mathbf{I}_{N_t} - \frac{\bar{\lambda}_k}{\gamma_k} \mathbf{A}^{-\frac{1}{2}} \mathbf{h}_{P_k} \mathbf{h}_{P_k}^H \mathbf{A}^{-\frac{H}{2}}$, we can express the left-hand side of (44b) as $\mathbf{A}^{\frac{1}{2}} \mathbf{B} \mathbf{A}^{\frac{H}{2}}$. Since \mathbf{A} is positive definite and \mathbf{B} is Hermitian, it is not difficult to show that $\mathbf{A}^{\frac{1}{2}} \mathbf{B} \mathbf{A}^{\frac{H}{2}} \succeq \mathbf{0}_{N_t}$ and

$\mathbf{B} \succeq \mathbf{0}_{N_t}$ are equivalent to each other. Then, the semi-definite positive constraints (44b) can be equivalently expressed as

$$\mathbf{I}_{N_t} - \frac{\bar{\lambda}_k}{\gamma_k} \left(\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j \mathbf{h}_{Pj} \mathbf{h}_{Pj}^H \right)^{-\frac{1}{2}} \mathbf{h}_{Pk} \mathbf{h}_{Pk}^H \left(\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j \mathbf{h}_{Pj} \mathbf{h}_{Pj}^H \right)^{-\frac{H}{2}} \succeq \mathbf{0}_{N_t}, \forall k. \quad (\text{D.1})$$

The positive semi-definite constraint in (D.1) means that the minimal eigenvalue of the matrix in the left-hand side should be non-negative. Note that the term $\left(\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j \mathbf{h}_{Pj} \mathbf{h}_{Pj}^H \right)^{-\frac{1}{2}}$ $\mathbf{h}_{Pk} \mathbf{h}_{Pk}^H \left(\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j \mathbf{h}_{Pj} \mathbf{h}_{Pj}^H \right)^{-\frac{H}{2}}$ is of rank one. It has only one positive eigenvalue, which is $\mathbf{h}_{Pk}^H \left(\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j \mathbf{h}_{Pj} \mathbf{h}_{Pj}^H \right)^{-1} \mathbf{h}_{Pk}$, and all other eigenvalues are zeros. Therefore, the constraint in (D.1) can be further converted into the constraint on the minimal eigenvalue of the matrix in the left-hand side of (D.1) as

$$1 - \frac{\bar{\lambda}_k}{\gamma_k} \mathbf{h}_{Pk}^H \left(\mathbf{I}_{N_t} + \sum_{j \neq k} \bar{\lambda}_j \mathbf{h}_{Pj} \mathbf{h}_{Pj}^H \right)^{-1} \mathbf{h}_{Pk} \geq 0, \forall k. \quad (\text{D.2})$$

which can be further rewritten as (45).

REFERENCES

- [1] B. Soret, H. Wang, K. I. Pedersen, and C. Rosa, "Multicell cooperation for LTE-advanced heterogeneous network scenarios," *IEEE Wireless Commun. Mag.*, vol. 20, no. 1, pp. 27–34, Feb. 2013.
- [2] C. Yang, S. Han, X. Hou, and A. F. Molisch, "How do we design CoMP to achieve its promised potential?" *IEEE Wireless Commun. Mag.*, vol. 20, no. 1, pp. 67–74, Feb. 2013.
- [3] J. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proc. ACM MOBICOM*, 2010.
- [4] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.
- [5] S. Hong, J. Brand, J. Choi, M. Jain, J. Mehlman, S. Katti, and P. Levis, "Applications of self-interference cancellation in 5G and beyond," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 114–121, Feb. 2014.
- [6] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 680–694, Feb. 2014.
- [7] T. Riihonen, S. Werner, and R. Wichman, "Optimized gain control for single-frequency relaying with loop interference," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2801–2806, June 2009.
- [8] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE J. Select. Areas Commun.*, vol. 30, no. 8, pp. 1541–1553, Sept. 2012.
- [9] G. Zheng, I. Krikidis, and B. Ottersten, "Full-duplex cooperative cognitive radio with transmit imperfections," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2498–2511, May 2013.
- [10] T. Riihonen, S. Werner, J. Cousseau, and R. Wichman, "Design of co-phasing allpass filters for full-duplex OFDM relays," in *Proc. ASILOMAR*, 2008.
- [11] T. Riihonen, R. Wichman, and J. Hamalainen, "Co-phasing full-duplex relay link with non-ideal feedback information," in *Proc. IEEE ISWCS*, 2008.
- [12] Y. Kishiyama, T. Uchino, S. Nagata, A. Morimoto, and Y. Xiang, "Heterogeneous network capacity expansion technology for further development of LTE/LTE-Advanced," *NTT DOCOMO Technical Journal*, vol. 15, no. 2, pp. 9–17, 2013.
- [13] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *IEEE Trans. Signal Processing*, vol. 60, no. 5, pp. 2508–2523, May 2012.
- [14] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2009.
- [15] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Processing*, vol. 55, pp. 2646–2660, June 2007.
- [16] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Processing*, vol. 54, no. 1, pp. 161–176, Jan. 2006.
- [17] C. D. Meyer, *Matrix analysis and applied linear algebra*. Siam, 2000.
- [18] T. L. Marzetta and B. M. Hochwald, "Fast transfer of channel state information in wireless systems," *IEEE Trans. Signal Processing*, vol. 54, no. 4, pp. 1268–1278, Apr. 2006.
- [19] 3GPP TR 36.814, "Further Advancements for E-UTRA Physical Layer Aspects (Release 9)," 2010.
- [20] M. Wildemeersch, T. Q. S. Quek, M. Kountouris, A. Rabbachin, and C. H. Slump, "Successive interference cancellation in heterogeneous networks," *IEEE Trans. Commun.*, vol. 62, no. 12, pp. 4440–4453, Dec. 2014.
- [21] R. Zakhour, Z. K. M. Ho, and D. Gesbert, "Distributed beamforming coordination in multicell MIMO channels," in *Proc. IEEE VTC Spring*, 2009.
- [22] 3GPP TS 36.211, "Evolved Universal Terrestrial Radio Access (E-UTRA): Physical channels and modulation," 2015.
- [23] IST-WINNER D1.1.2 P. Kyösti, et al., "WINNER II Channel Models v1.2," Sep. 2007.